

# **PubChem: An Open Repository for Chemical Structure and Biological Activity Information**

**Stephen H. Bryant**

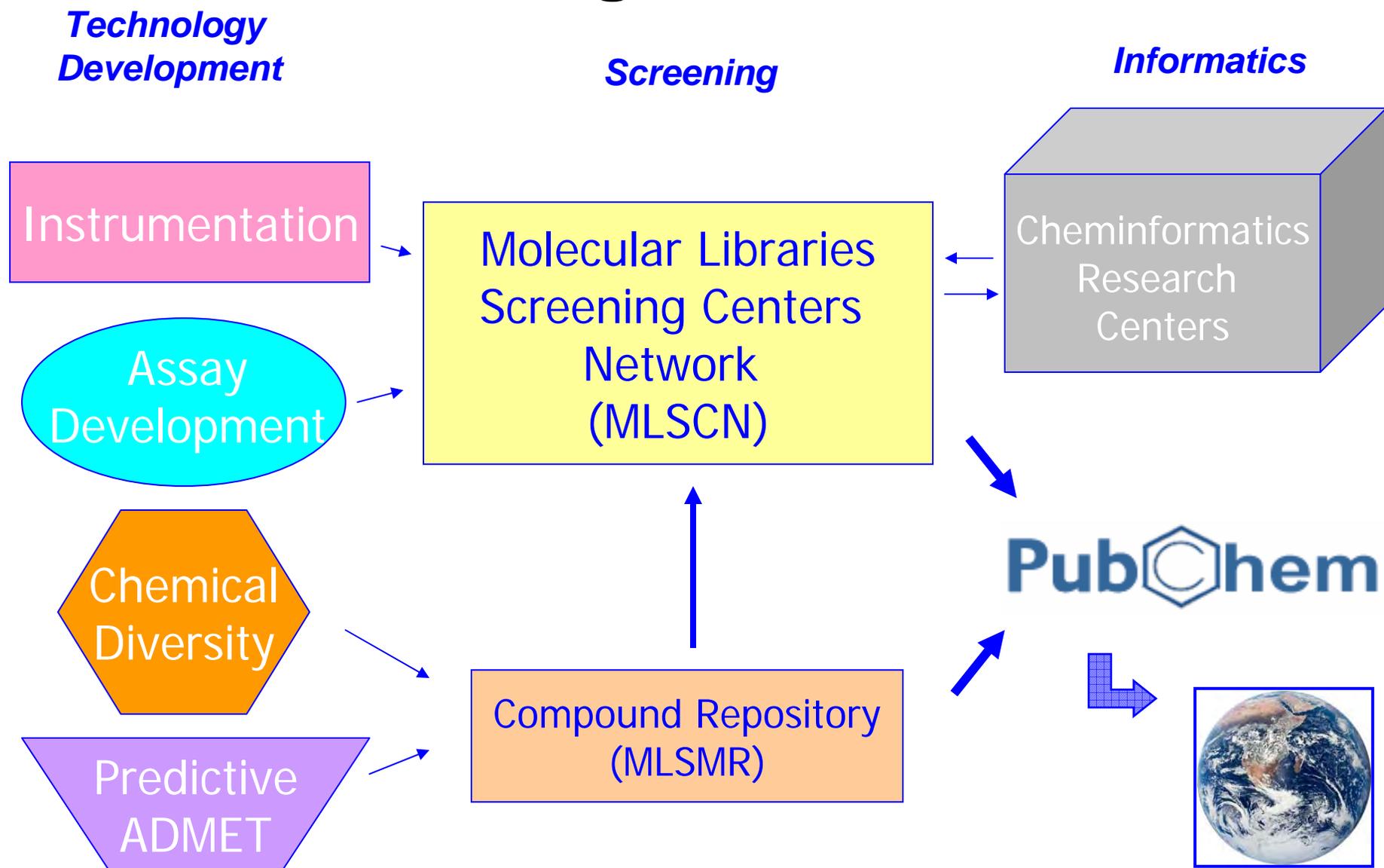
**Forum on Computational  
Toxicology**

**Research Triangle Park**

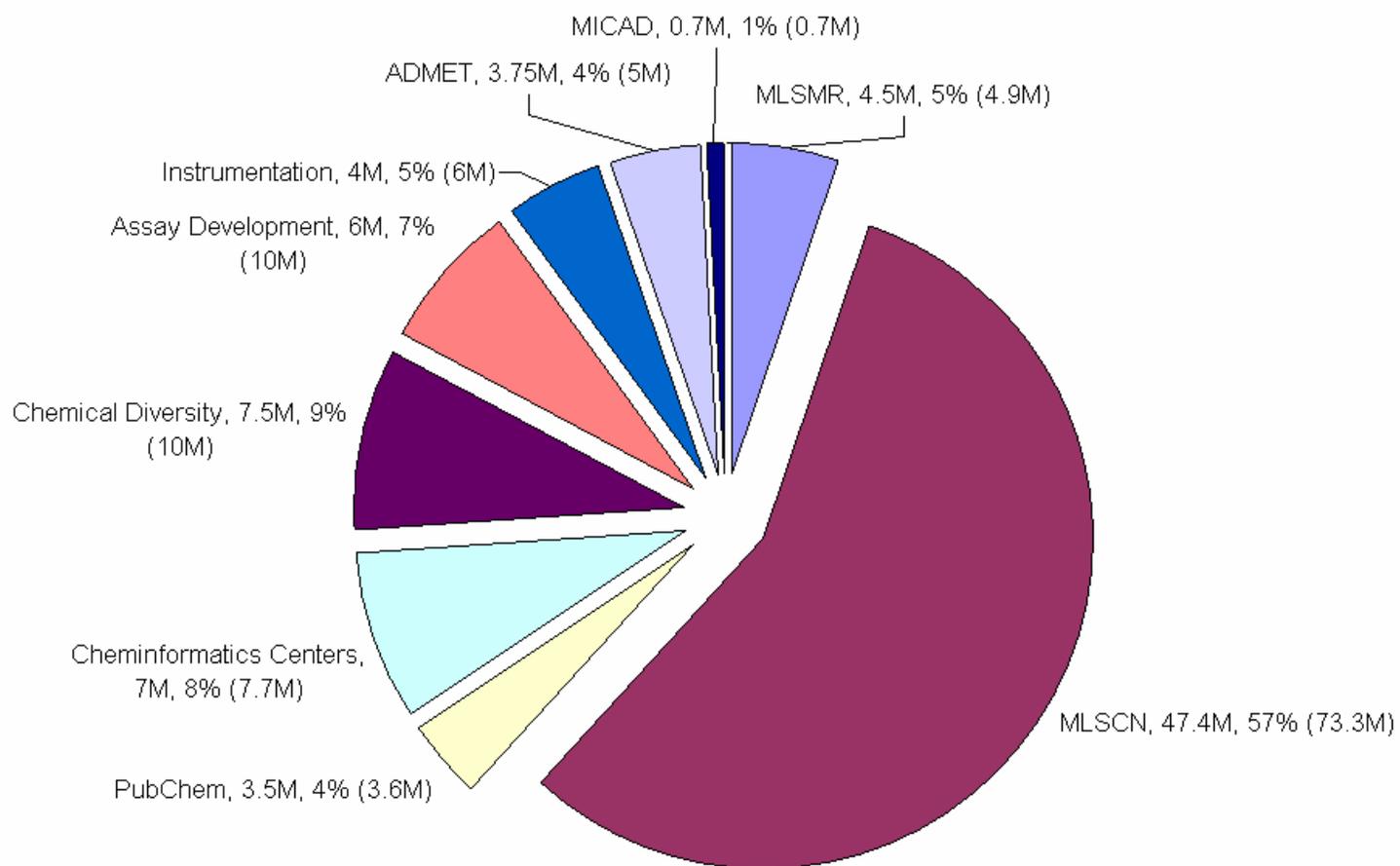
**May 21-23, 2007**



# The Molecular Libraries Program: An Integrated Initiative



# Molecular Libraries Components ...



# PubChem Goals ...

- ... Archive molecular structure and bioassay data from the Molecular Libraries Screening Center Network**
- ... Provide search, retrieval and data analysis tools to optimize utility**

# PubChem Goals ...

**... Further optimize research utility by including other public sources of chemical structure and bioactivity information**

**... and by integration with other NIH Biomedical information resources whenever possible**

# **PubChem Approach ...**

**... “GenBank model”**

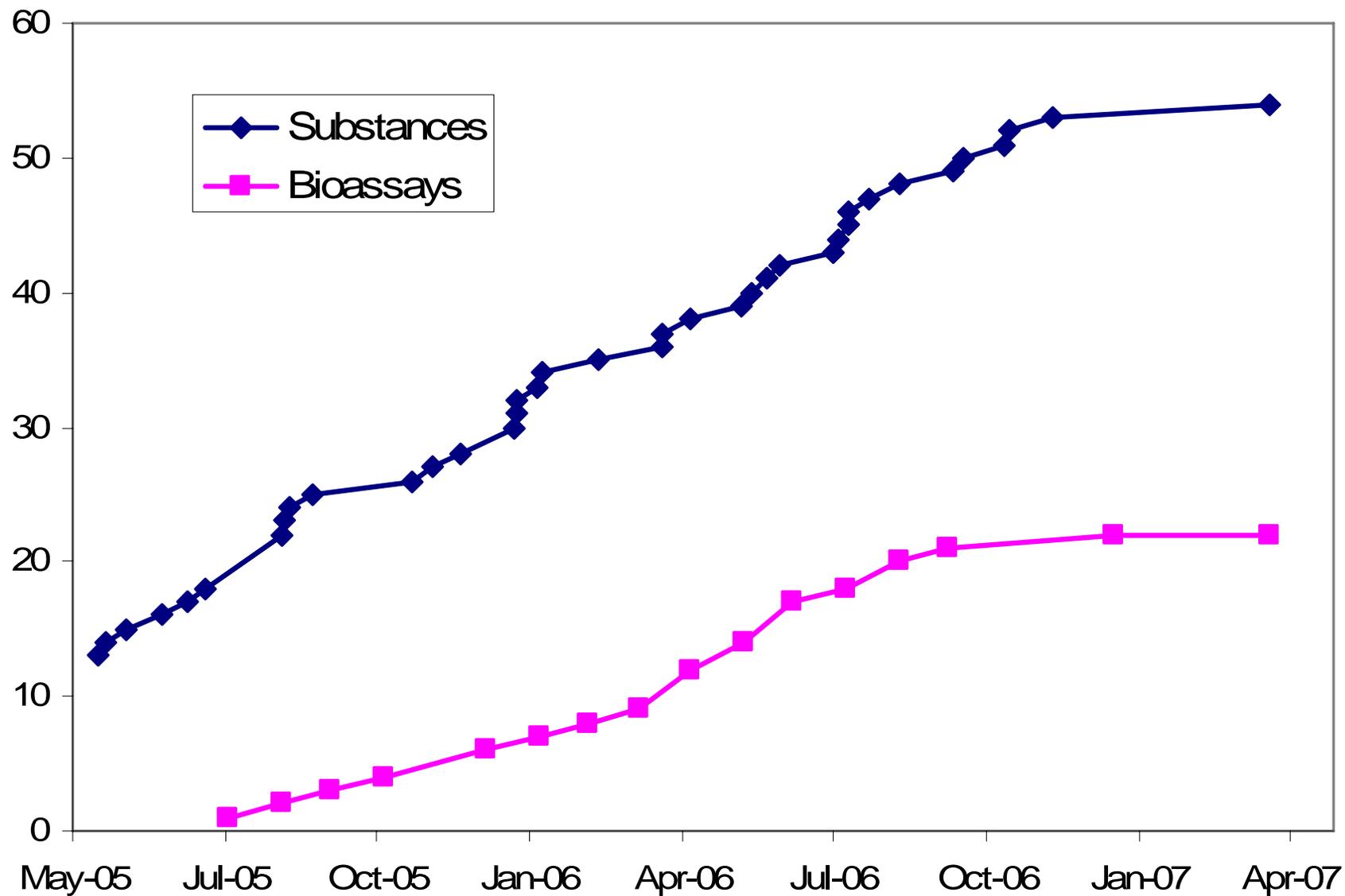
**... direct depositions by investigators**

**... highly automated (low database cost)**

**... 25 year precedents in biology**

**... less precedent in chemistry**

# Growth In PubChem Contributing Organizations



# **PubChem Contents ...**

**... Contributed substance records**

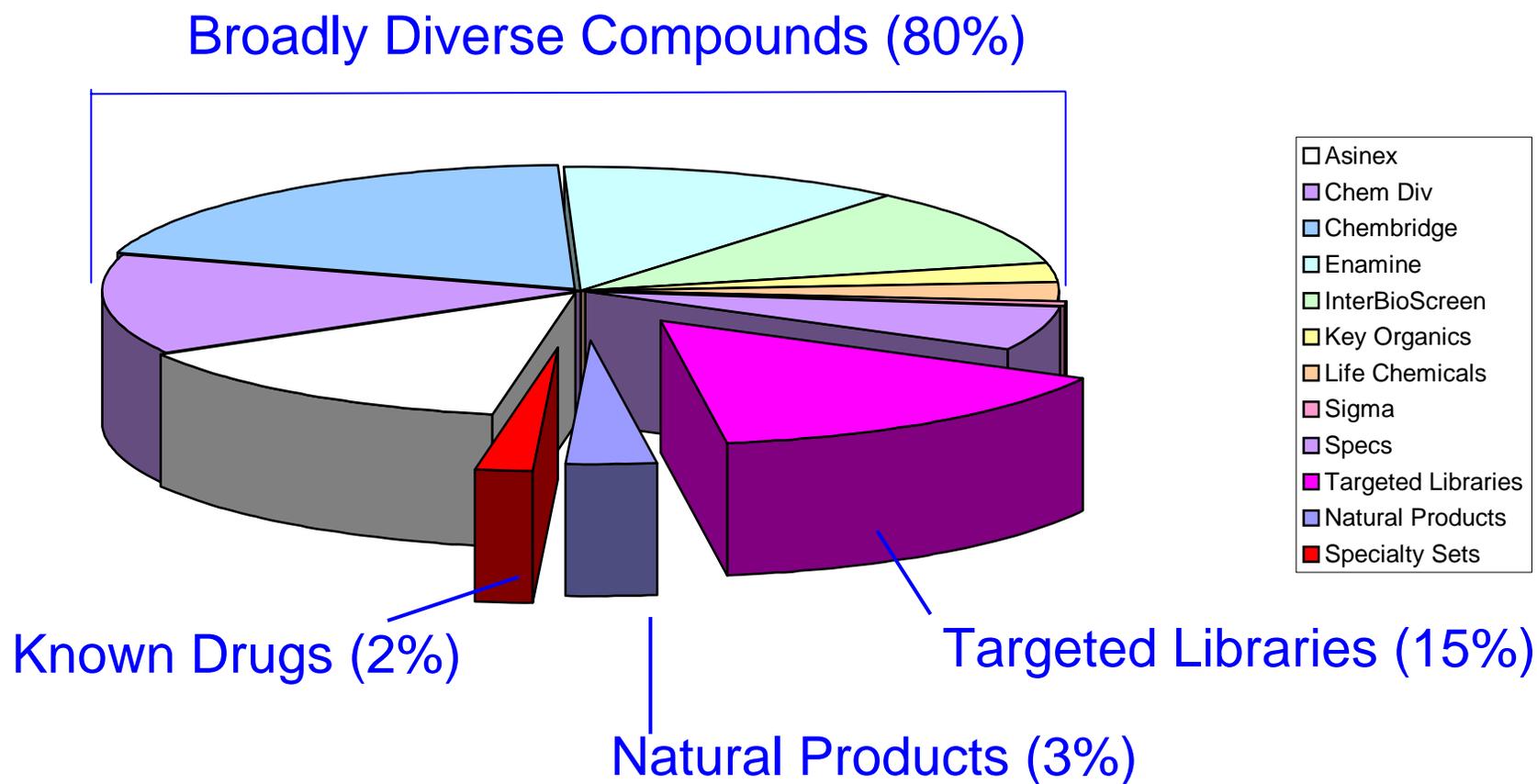
**... with chemical structure**

**... chemical names and comments**

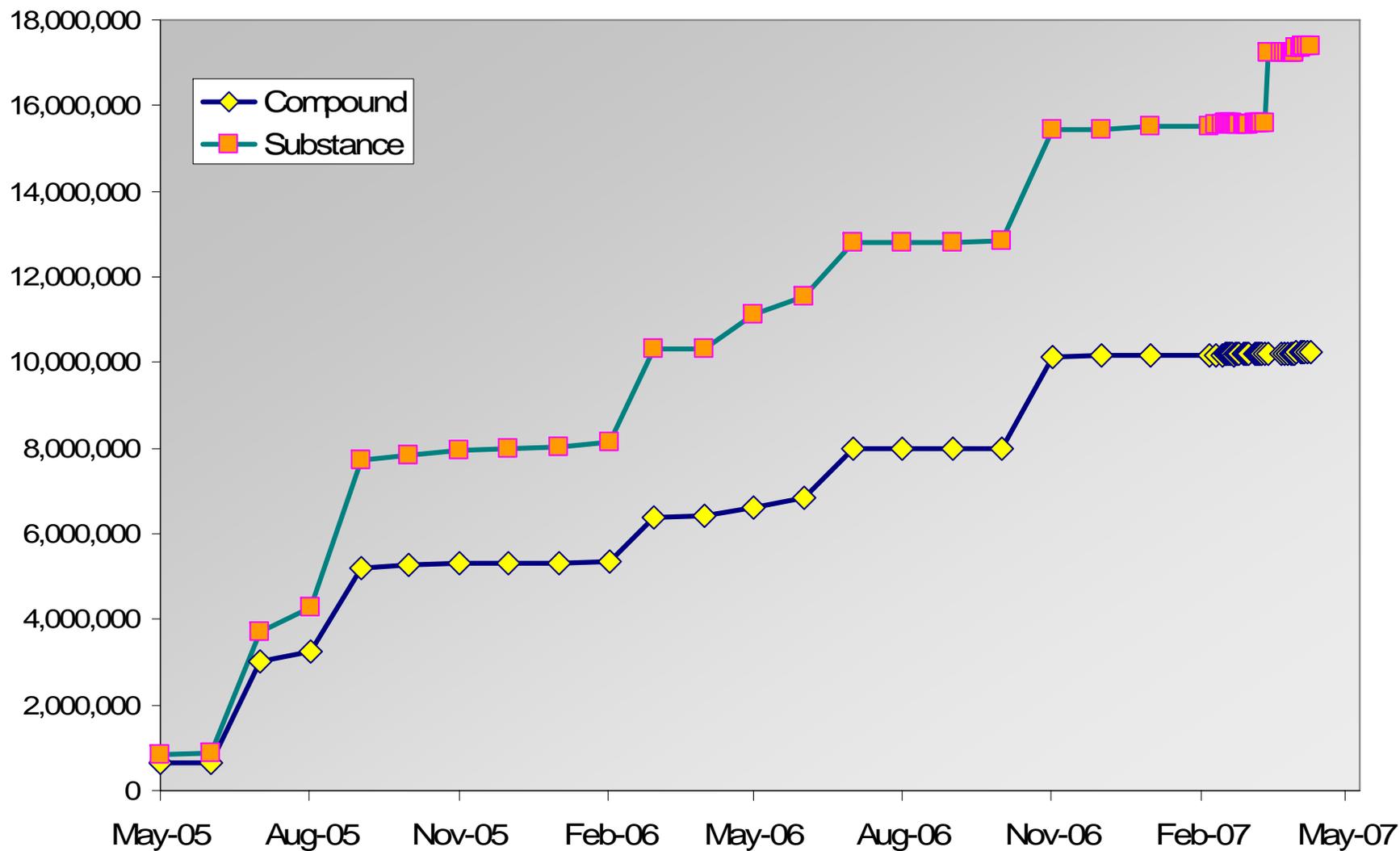
**... links to contributor web sites**

**... contributed links to other NCBI  
biomedical databases**

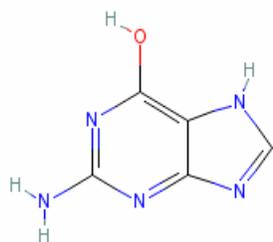
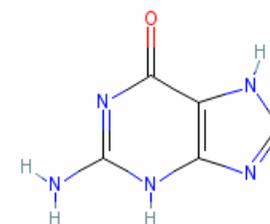
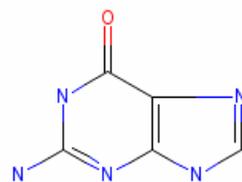
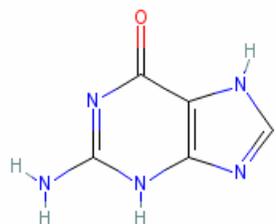
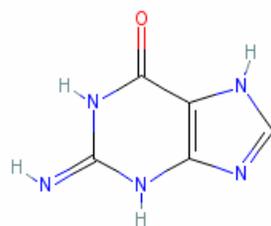
# To Date 203,554 Molecular Libraries Compounds in PubChem



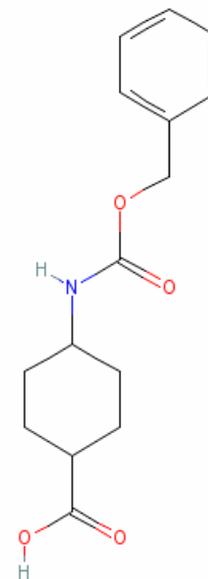
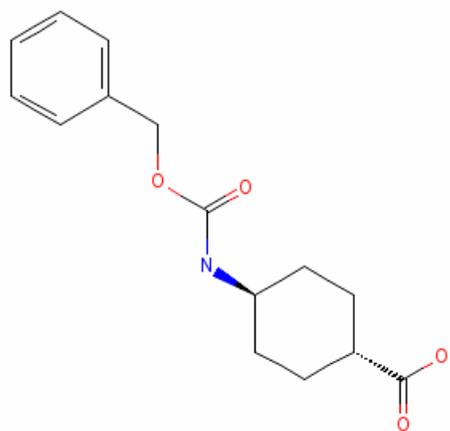
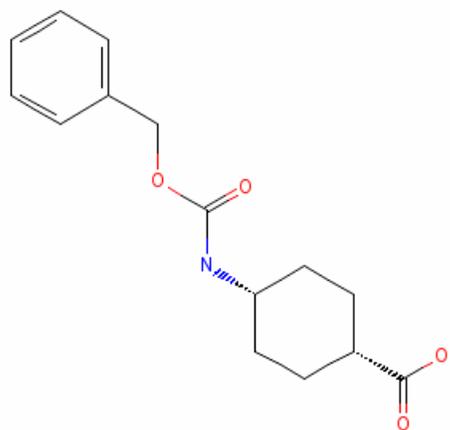
# Growth In PubChem Substances / Compounds



# Compound Challenges ...



# PubChem Standardization ...



# **PubChem Contents ...**

**... Contributed bioassay records**

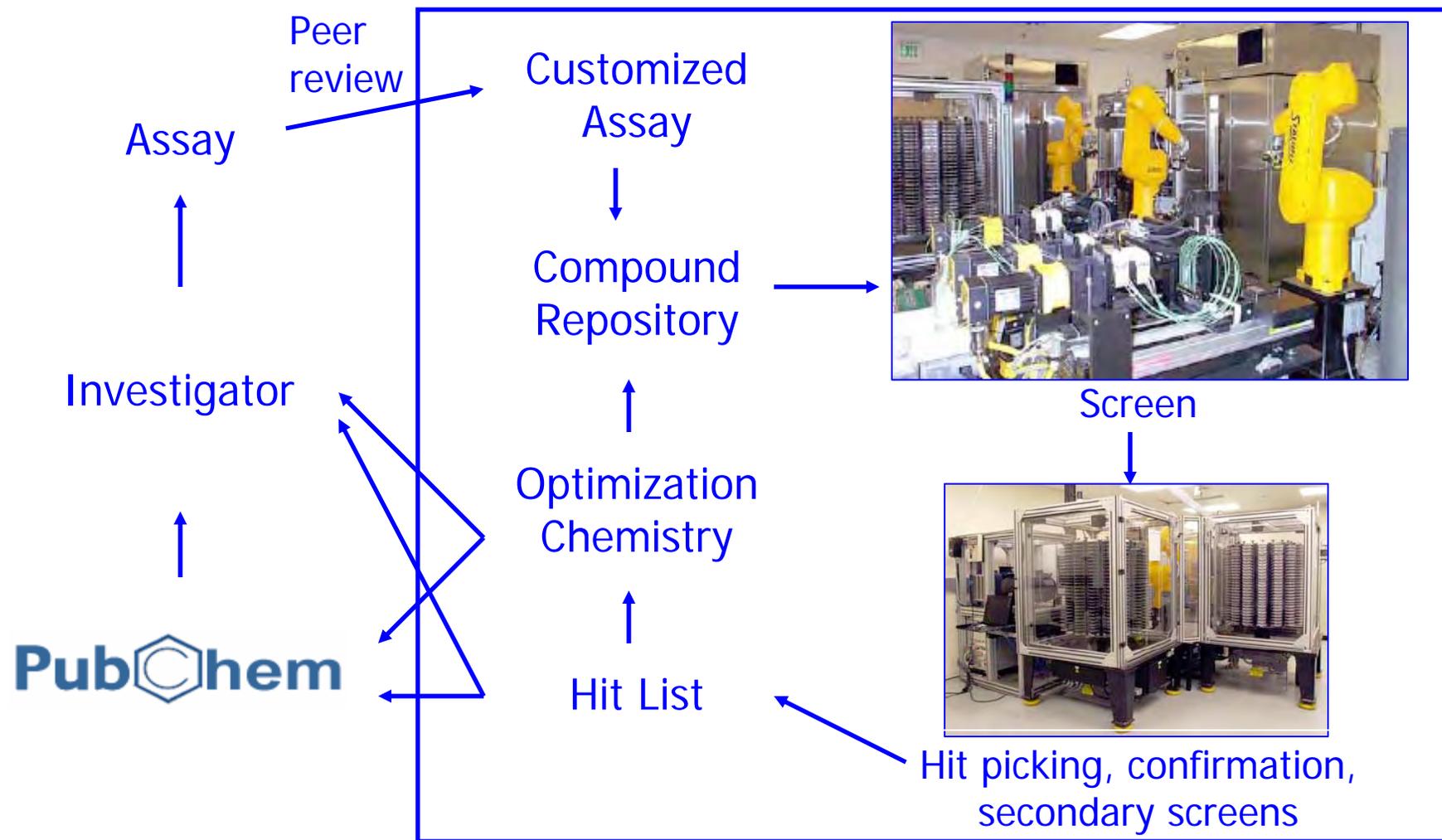
**... with assay description / protocol**

**... links to tested substances**

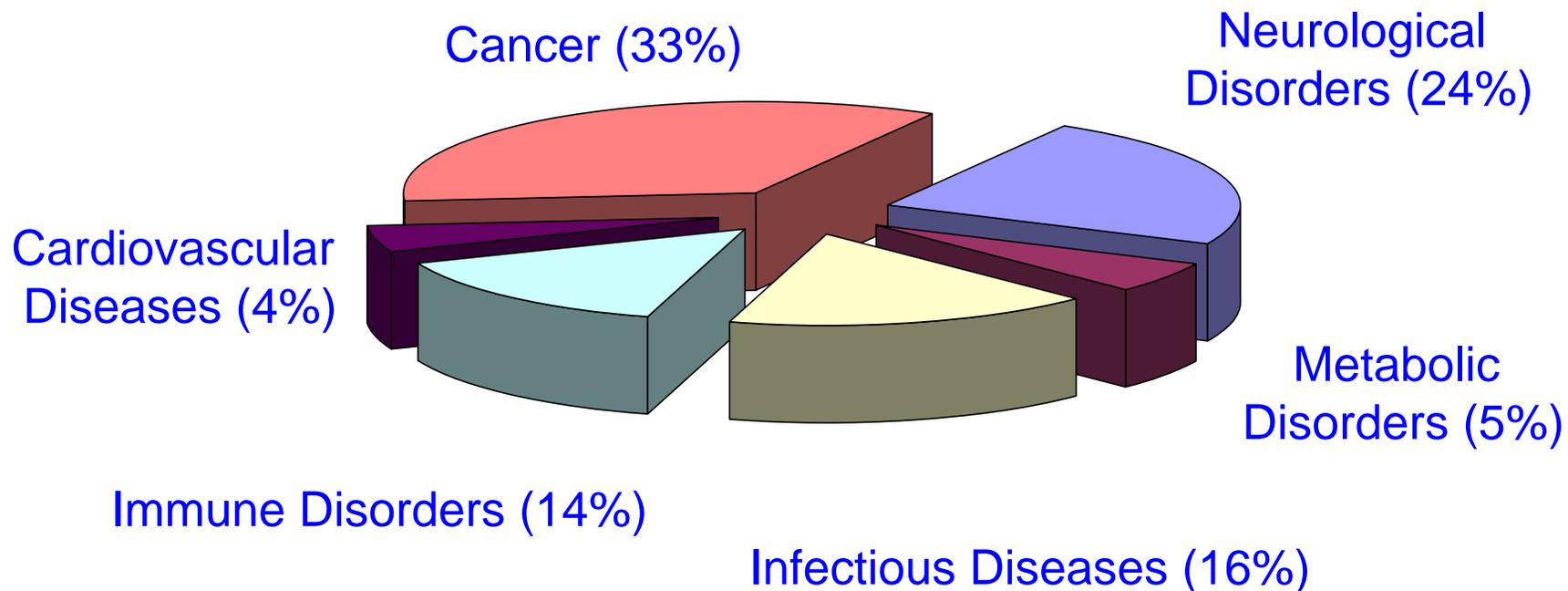
**... summary and detailed test results**

**... links to contributor web sites and other  
NCBI databases**

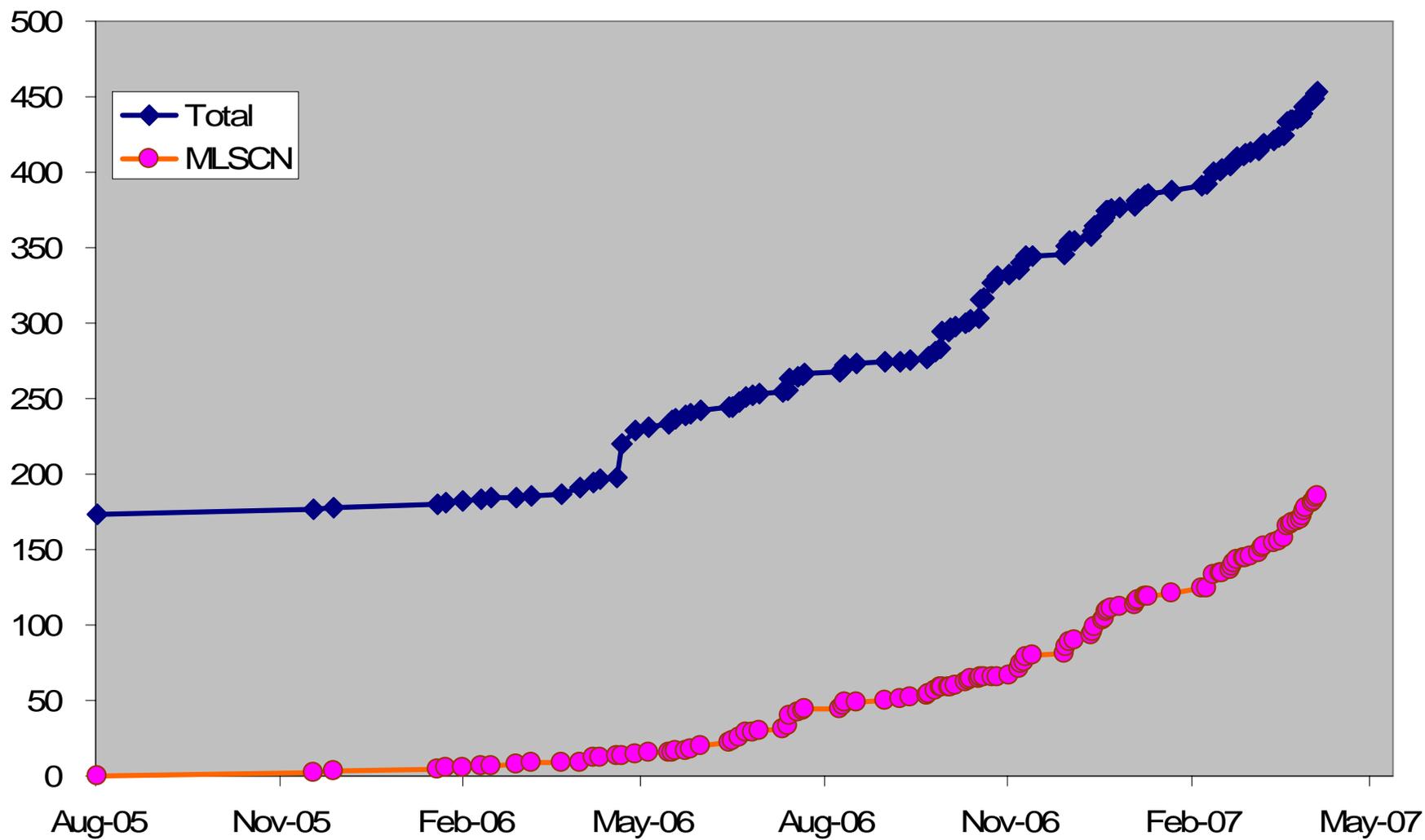
# Molecular Libraries BioAssays ...



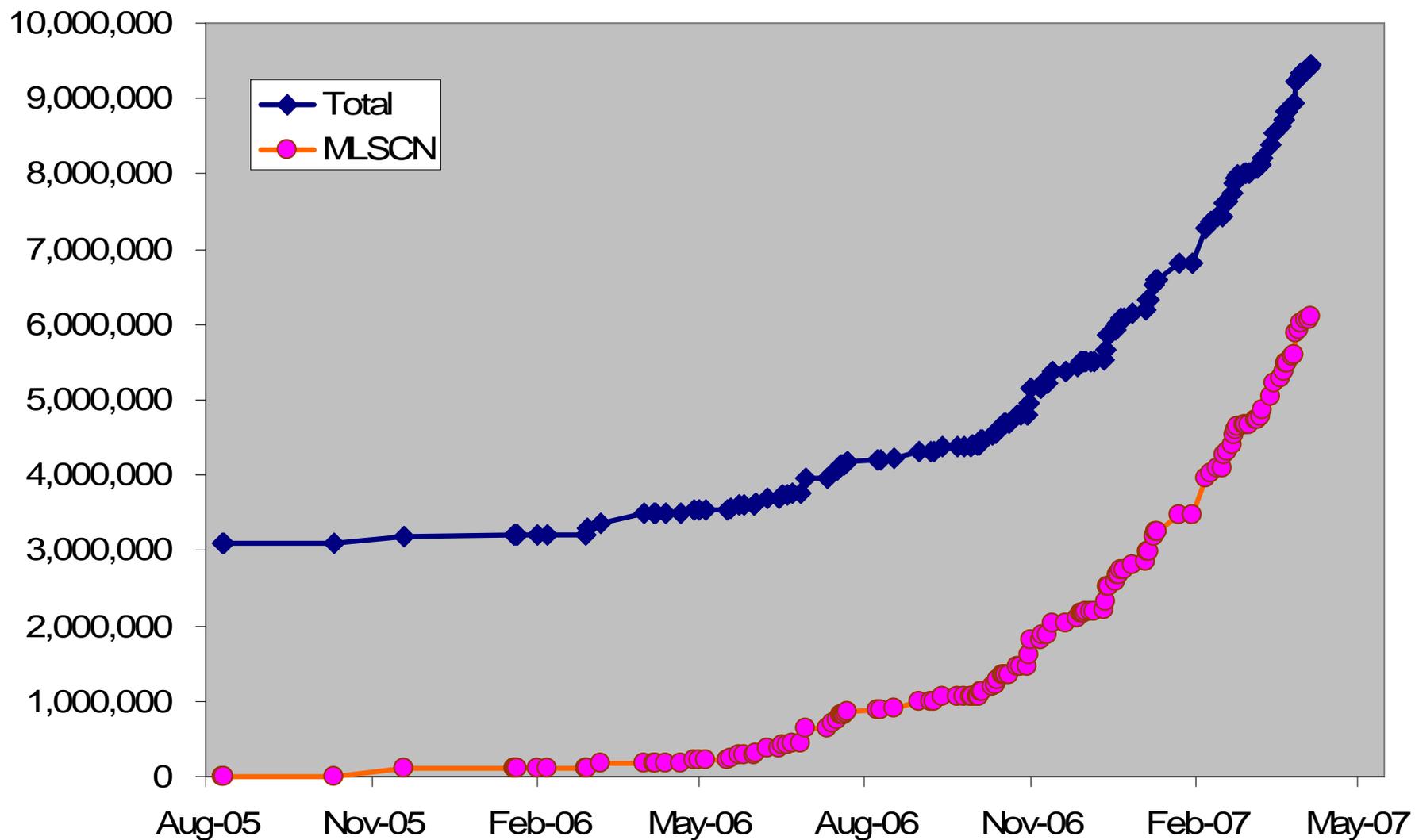
# To Date 248 Molecular Libraries BioAssays in PubChem



# Growth In PubChem BioAssays



# Growth In PubChem Tested Substances



# EPA “DSSTox” BioAssays ...

The screenshot shows a Windows Internet Explorer browser window displaying the PubChem BioAssay search results page. The address bar shows the URL <http://www.ncbi.nlm.nih.gov/sites/entrez>. The search query is `epa_dsstox[sourcename]` in the PubChem BioAssay database. The results are displayed in a list format, showing two items:

- Item 1:** AID: [355](#)  
Water Disinfection By-Products with Predicted Carcinogenicity Estimates (DBPCAN)  
Source: EPA DSSTox  
Total substance tested: 209; Active: 0
- Item 2:** AID: [356](#)  
EPA Fathead Minnow Acute Toxicity database (EPAFHM)  
Source: EPA DSSTox  
Total substance tested: 617; Active: 0

The interface includes navigation tabs for Limits, Preview/Index, History, Clipboard, and Details. The display settings are set to Summary, showing 20 items per page. The status bar at the bottom indicates 'Done' and 'Internet' connectivity.

# NCGC / NTP / NIEHS Screens ...

The screenshot shows a Windows Internet Explorer browser window displaying the NCBI PubChem BioAssay search results for the query "ntp niehs". The browser's address bar shows the URL "http://www.ncbi.nlm.nih.gov/sites/entrez". The page header includes the NCBI logo, the PubChem BioAssay logo, and the National Library of Medicine (NLM) logo. A navigation bar at the top lists various databases: All Databases, PubMed, Nucleotide, Protein, Genome, Structure, PMC, PubChem, and Books. The search bar contains the text "Search PubChem BioAssay for ntp niehs" with "Go", "Clear", and "Save Search" buttons. Below the search bar, there are tabs for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The "Display" dropdown is set to "Summary", and the "Show" dropdown is set to "20". The search results are displayed in a list format, showing two items:

- 1: AID: [426](#) Links  
Cell Viability - Jurkat  
Source: NCGC  
Total substance tested: 2816; Active: 284
- 2: AID: [427](#) Links  
Cell Viability - Hek293  
Source: NCGC  
Total substance tested: 2816; Active: 160

The page also includes a sidebar with links to "About Entrez", "Entrez Help", "PubChem Help | FAQ", "PubChem Substance Structures supplied by depositors", and "PubChem Compound". The bottom of the browser window shows the "Internet" icon and a "100%" zoom level.

# BioAssay Challenges ...

The screenshot shows a Windows Internet Explorer browser window displaying the PubChem BioAssay Summary for AID: 706. The browser's address bar shows the URL <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=706>. The page title is "BioAssay Summary".

**BioAssay Summary**

**AID: 706**  
**Name:** [Yeast Lifespan Shortening Chemical Screening, Restrictive Growth Control - Pilot Screen](#)  
**Data Source:** [SRMLSC \(Yeast Lifespan Pilot - restrictive\)](#)

**Test Results:** [Show](#) [Select](#) [Plot](#)

**Structure Activity Analysis:** [Show](#)

**Structure Clustering:** [Show](#)

**Neighbors, Related BioAssays:**

- Activity Overlap:** [81](#) [Summary](#)
- Compound BioActivity Summary:** [Active](#) [All](#)
- Substance BioActivity Summary:** [Active](#) [All](#)

The browser's taskbar at the bottom shows the "Internet" icon and a zoom level of 100%.

# BioAssays are Diverse ...

The screenshot shows a Windows Internet Explorer browser window displaying the PubChem BioAssay Summary for AID 679. The browser's address bar shows the URL <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=679>. The page title is "BioAssay Summary".

The main content area displays the following information:

- AID: 679** (with a document icon)
- Name:** [Factor XIa Dose Response Confirmation](#)
- Data Source:** PCMD (FXIA\_DRCONF)

Below this information are three interactive sections:

- Test Results:** Includes buttons for [Show](#), [Select](#), and [Plot](#).
- Structure Activity Analysis:** Includes a [Show](#) button.
- Structure Clustering:** Includes a [Show](#) button.

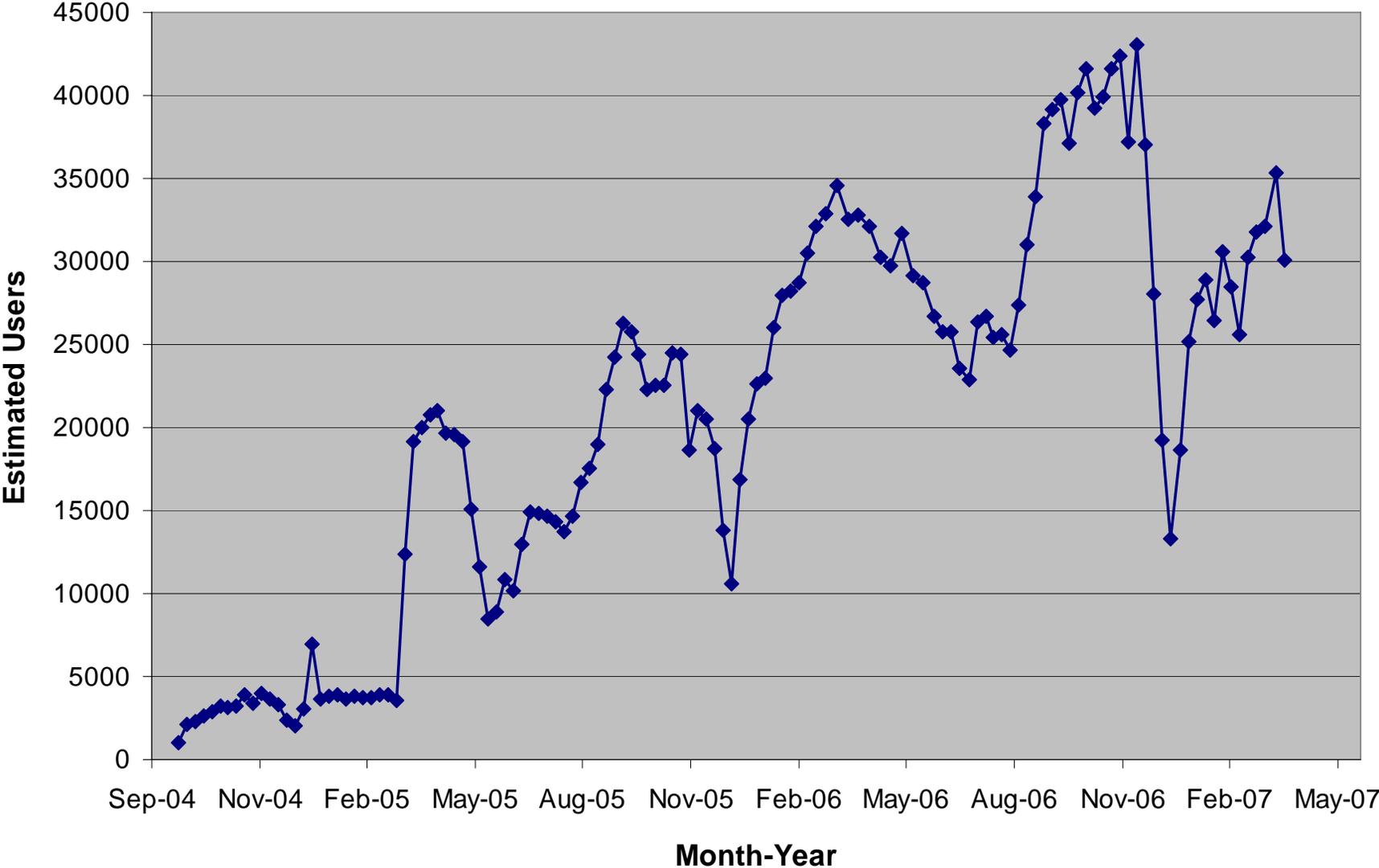
A horizontal navigation bar contains the following tabs: [Neighbors](#), [Links](#), [Description](#), [Protocol](#), [Comment](#), and [Definitions](#).

The **Neighbors, Related BioAssays:** section (with a magnifying glass icon) lists the following related items:

- Activity Overlap:** 34 [Summary](#)
- Target Similarity:** 6 [Summary](#)
- Compound BioActivity Summary:** [Active](#) [All](#)
- Substance BioActivity Summary:** [Active](#) [All](#)

The browser's status bar at the bottom shows "Internet" and a zoom level of "100%".

# Engineering Challenges ...



# PubChem Retrieval System ...

**... Optimize “discoverability” for molecular biologists by integrating PubChem into NCBI’s Entrez / PubMed Search Engine**

... Chemical structure search

... Bioassay result search

... Exploratory structure-activity tools

# NCBI's Entrez Search Engine ...

The screenshot shows the NCBI Entrez cross-database search page in Microsoft Internet Explorer. The browser window title is "Entrez cross-database search - Microsoft Internet Explorer". The address bar shows the URL "http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi". The page features the NCBI logo and the Entrez logo with the tagline "Entrez, The Life Sciences Search Engine". A navigation menu includes links for HOME, SEARCH, SITE MAP, PubMed, All Databases, Human Genome, GenBank, Map Viewer, and BLAST. A search bar is present with "GO" and "CLEAR" buttons. The main content area is titled "Welcome to the Entrez cross-database search page" and lists various databases with their respective icons and descriptions. Each database entry includes a question mark icon for help.

NCBI

Entrez, The Life Sciences Search Engine

HOME | SEARCH | SITE MAP | PubMed | All Databases | Human Genome | GenBank | Map Viewer | BLAST

Search across databases  GO CLEAR Help

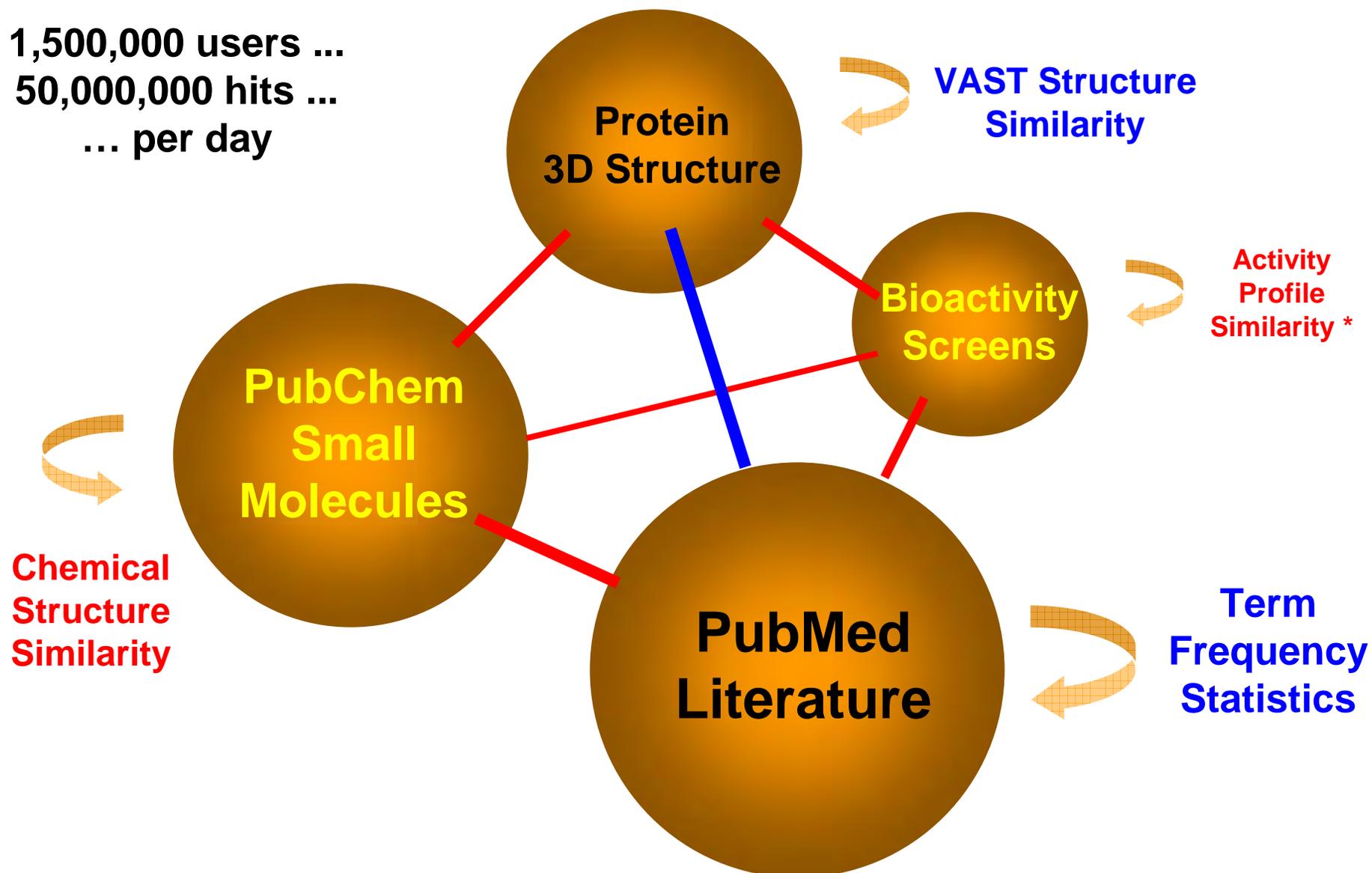
Welcome to the Entrez cross-database search page

<b>PubMed:</b> biomedical literature citations and abstracts	<b>Books:</b> online books
<b>PubMed Central:</b> free, full text journal articles	<b>OMIM:</b> online Mendelian Inheritance in Man
	<b>Site Search:</b> NCBI web and FTP sites
<b>Nucleotide:</b> sequence database (GenBank)	<b>UniGene:</b> gene-oriented clusters of transcript sequences
<b>Protein:</b> sequence database	<b>CDD:</b> conserved protein domain database
<b>Genome:</b> whole genome sequences	<b>3D Domains:</b> domains from Entrez Structure
<b>Structure:</b> three-dimensional macromolecular structures	<b>UniSTS:</b> markers and mapping data
<b>Taxonomy:</b> organisms in GenBank	<b>PopSet:</b> population study data sets
<b>SNP:</b> single nucleotide polymorphism	<b>GEO Profiles:</b> expression and molecular abundance profiles
<b>Gene:</b> gene-centered information	<b>GEO DataSets:</b> experimental sets of GEO data

Internet

# Entrez Links and Neighbors ...

1,500,000 users ...  
50,000,000 hits ...  
... per day



# Search for “Shoichet inhibitors” ...

Entrez cross-database search - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi

Entrez cross-database search

NCBI Entrez, The Life Sciences Search Engine

HOME SEARCH SITE MAP PubMed All Databases Human Genome GenBank Map Viewer BLAST

Search across databases  GO CLEAR Help

51	PubMed: biomedical literature citations and abstracts	?	none	Books: online books	?	
33	PubMed Central: free, full text journal articles	?	none	OMIM: online Mendelian Inheritance in Man	?	
10	Site Search: NCBI web and FTP sites	?	none	OMIA: Online Mendelian Inheritance in Animals	?	
3303	Nucleotide: sequence database (includes GenBank)	?		34	UniGene: gene-oriented clusters of transcript sequences	?
35	Protein: sequence database	?		112	CDD: conserved protein domain database	?
115	Genome: whole genome sequences	?		41	3D Domains: domains from Entrez Structure	?
11	Structure: three-dimensional macromolecular structures	?		26	UniSTS: markers and mapping data	?
none	Taxonomy: organisms in GenBank	?		63	PopSet: population study data sets	?
none	SNP: single nucleotide polymorphism	?		234943	GED Profiles: expression and molecular abundance profiles	?

Done Internet 100%

# PubMed Article Retrieved ...

Entrez PubMed - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=... Google

Entrez PubMed

NCBI PubMed  
A service of the National Library of Medicine and the National Institutes of Health  
www.pubmed.gov

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search PubMed for [ ] Go Clear

Limits Preview/Index History Clipboard Details

Display AbstractPlus Show 20 Sort by Send to

All: 1 Review: 0

1: [Nat Chem Biol. 2005 Aug;1\(3\):146-8. Epub 2005 Jul 3.](#)

Comment in:  
[Nat Chem Biol. 2005 Aug;1\(3\):125.](#)

**High-throughput assays for promiscuous inhibitors.**

[Feng BY](#), [Shelat A](#), [Doman TN](#), [Guy RK](#), [Shoichet BK](#).

Department of Pharmaceutical Chemistry & Graduate Group in Chemistry and Chemical Biology, 1700 4th St., University of California San Francisco, San Francisco, California 94143-2550, USA.

High-throughput screening (HTS) searches large libraries of chemical...

Related Links

- ▶ A common med promiscuous in [J Med Chem. 2002]
- ▶ Identification and prediction of promiscuous aggre [J Med Chem. 2003]
- ▶ Virtual screening to enrich hit lists from high-throughput screer [Proteins. 2003]
- ▶ Kinase inhibitors: not just for kinases anymore. [J Med Chem. 2003]

Links

- ▶ BioAssay
- ▶ Substance via MeSH
- ▶ Substance (Publisher)
- ▶ Cited in PMC
- ▶ LinkOut

Internet 100%

# Link to PubChem Records ...

PubChem Substance - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?tool=pubmed\_DocSum&db=pubm Google

PubChem Substance

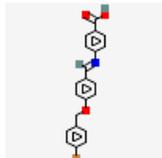
Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

All: 29 BioAssay: 0 Protein3D: 0 Rule of 5: 24

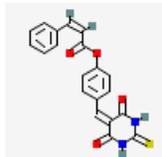
Items 1 - 20 of 29 Page 1 of 2 Next

1: SID: [855510](#) [Related Structures, Other Links](#)



CID: [657179](#), nchembio718-comp29, (E)-4-(4-(4-bromobenzyloxy)benzylideneamino)benzoic acid  
Source: [Nature Chemical Biology\(nchembio718-comp29\)](#)  
IUPAC: 4-[[4-[(4-bromophenyl)methoxy]phenyl]methylideneamino]benzoic acid  
MW: 410.261 | MF: C21H16BrNO3

2: SID: [855509](#) [Related Structures, Other Links](#)



CID: [5388956](#), nchembio718-comp28, (Z)-4-((4,6-dioxo-2-thioxo-tetrahydropyrimidin-5(6H)-ylidene)methyl)phenyl 3-phenylacrylate  
Source: [Nature Chemical Biology\(nchembio718-comp28\)](#)  
IUPAC: [4-[(4,6-dioxo-2-sulfanylidene-1,3-diazinan-5-ylidene)methyl]phenyl] (Z)-3-phenylprop-2-enoate  
MW: 378.402 | MF: C20H14N2O4S

PubChem Substance Structures supplied by depositors

PubChem Compound Unique structures with computed properties

PubChem BioAssay

Internet 100%

# “Kaempferol” in PubChem ...

CID 576 -- PubChem Compound Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Copy Paste

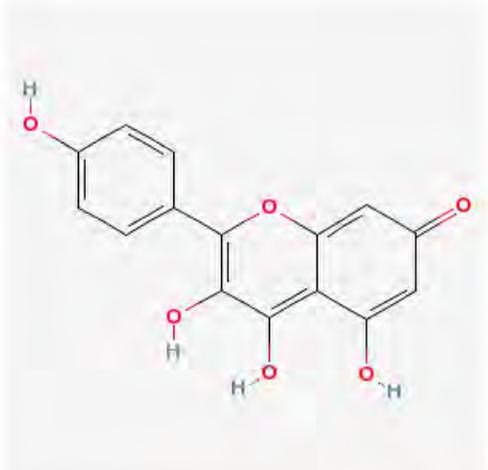
Address <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=576> Go Links »

NCBI PubChem National Library of Medicine NLM

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

Search PubChem Compound  GO

### Compound Summary:



Chemical structure of Kaempferol (CID: 576) is displayed. The structure shows a flavone core with a 4-hydroxyphenyl group at position 7, and hydroxyl groups at positions 3, 5, and 8. The SMILES string is Oc1ccc(cc1)-c2c(O)c(O)c(O)c2=O.

-  CID: 576 [?](#)
-  Substances: [?](#)
  - All: 14 Links
  - Same: 12 Links
  - Mixture: 2 Links
-  BioActivity: 7 Links [?](#)
-  PubMed: 16 Links [?](#)
-  Protein Structure: 1 Link [?](#)
-  NLM Toxicology: [Link](#) [?](#)
-  Similar Compounds: 73 Links [?](#)
-  Structure Search [?](#)

MeSH Synonyms Properties Descriptors Category Exports

Internet

# Similar Compounds in PubChem ...

PubChem Compound - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=DisplayFiltered&DB=pccomp

PubChem Compound

ABOUT ENTREZ  
Entrez Help

PubChem  
Help | FAQ

PubChem  
Substance  
Structures  
supplied by  
depositors

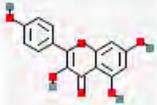
PubChem  
Compound  
Unique structures  
with computed  
properties

PubChem  
BioAssay  
Bioactivity assay  
results supplied by  
depositors

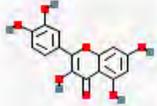
All: 1481 BioAssay: 163 Protein3D: 12 Rule of 5: 1257

Items 1 - 12 of 12 One page.

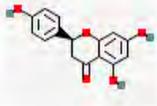
1: CID: [5280863](#) Related Structures, Assays, Literature, Other Entrez Databases

 kaempferol, Kaempherol ...  
IUPAC: 3,5,7-trihydroxy-2-(4-hydroxyphenyl)chromen-4-one  
MW: 286.236 | MF: C15H10O6

2: CID: [5280343](#) Related Structures, Assays, Literature, Other Entrez Databases

 quercetin, Sophoretin ...  
IUPAC: 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-chromen-4-one  
MW: 302.236 | MF: C15H10O7

3: CID: [439246](#) Related Structures, Assays, Literature, Other Entrez Databases

 naringenin, naringetol ...  
IUPAC: (2S)-5,7-dihydroxy-2-(4-hydroxyphenyl)chroman-4-one  
MW: 272.253 | MF: C15H12O5

Done Internet 100%

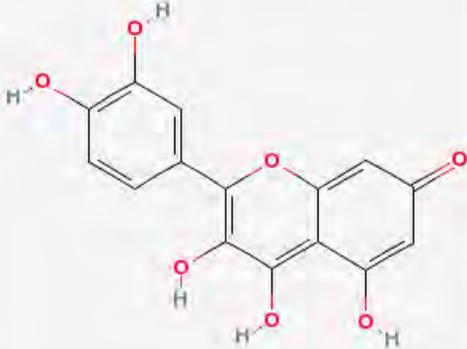
# “Quercetin” in PubChem ...

CID 480 -- PubChem Compound Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Download Upload

Address <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=480> Go Links »



**CID: 480** ⓘ

**Substances:** ⓘ  
All: 21 Links  
Same: 16 Links  
Mixture: 5 Links

**BioActivity:** 74 Links ⓘ

**PubMed:** 27 Links ⓘ

**Protein Structures:** 4 Links ⓘ

**NLM Toxicology:** Link ⓘ

**Related Compounds:** ⓘ  
Same, Any Tautomers: 2 Links

**Similar Compounds:** 109 Links ⓘ

**Structure Search** ⓘ

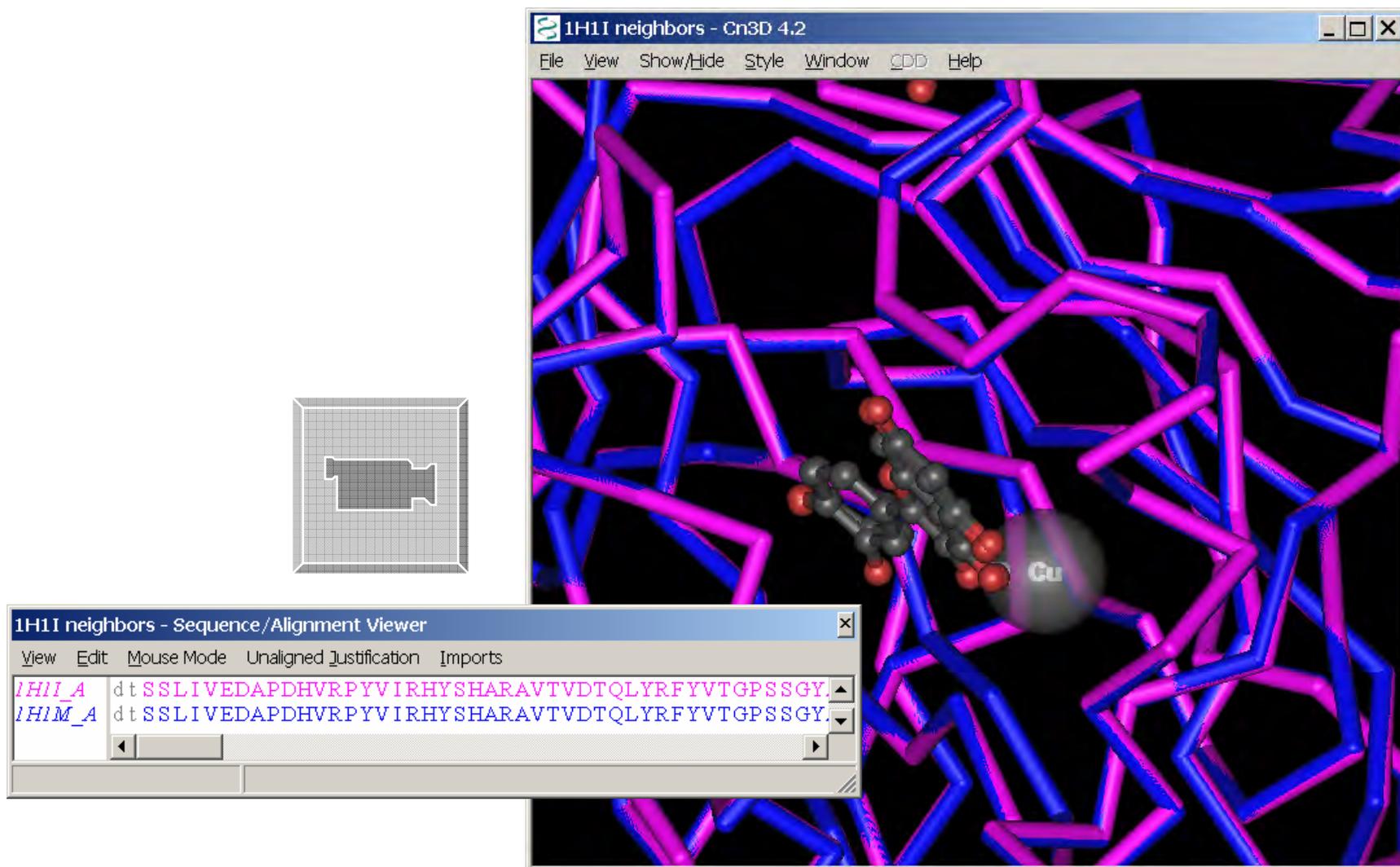
MeSH Synonyms Properties Descriptors Category Exports

**Medical Subject Annotations:** (Total:2) ⓘ Display: Next 1 | All

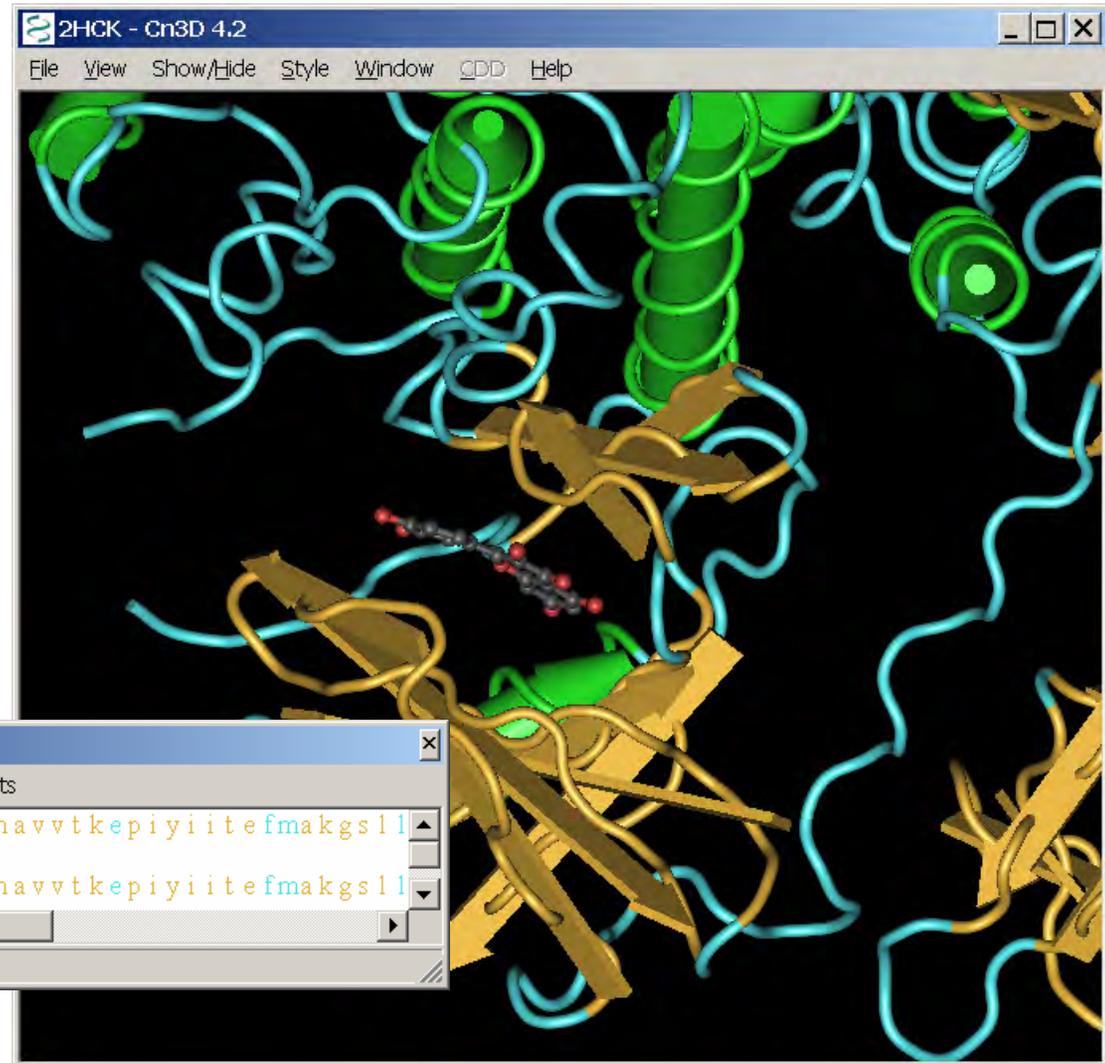
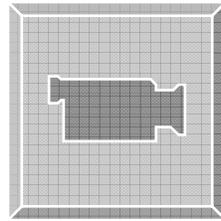
**Quercetin**  
A flavonol widely distributed in plants. It is an antioxidant. Like many other phenolic

Internet

# Compare Protein / Ligand Complexes ...



# Link to Another Structure ...

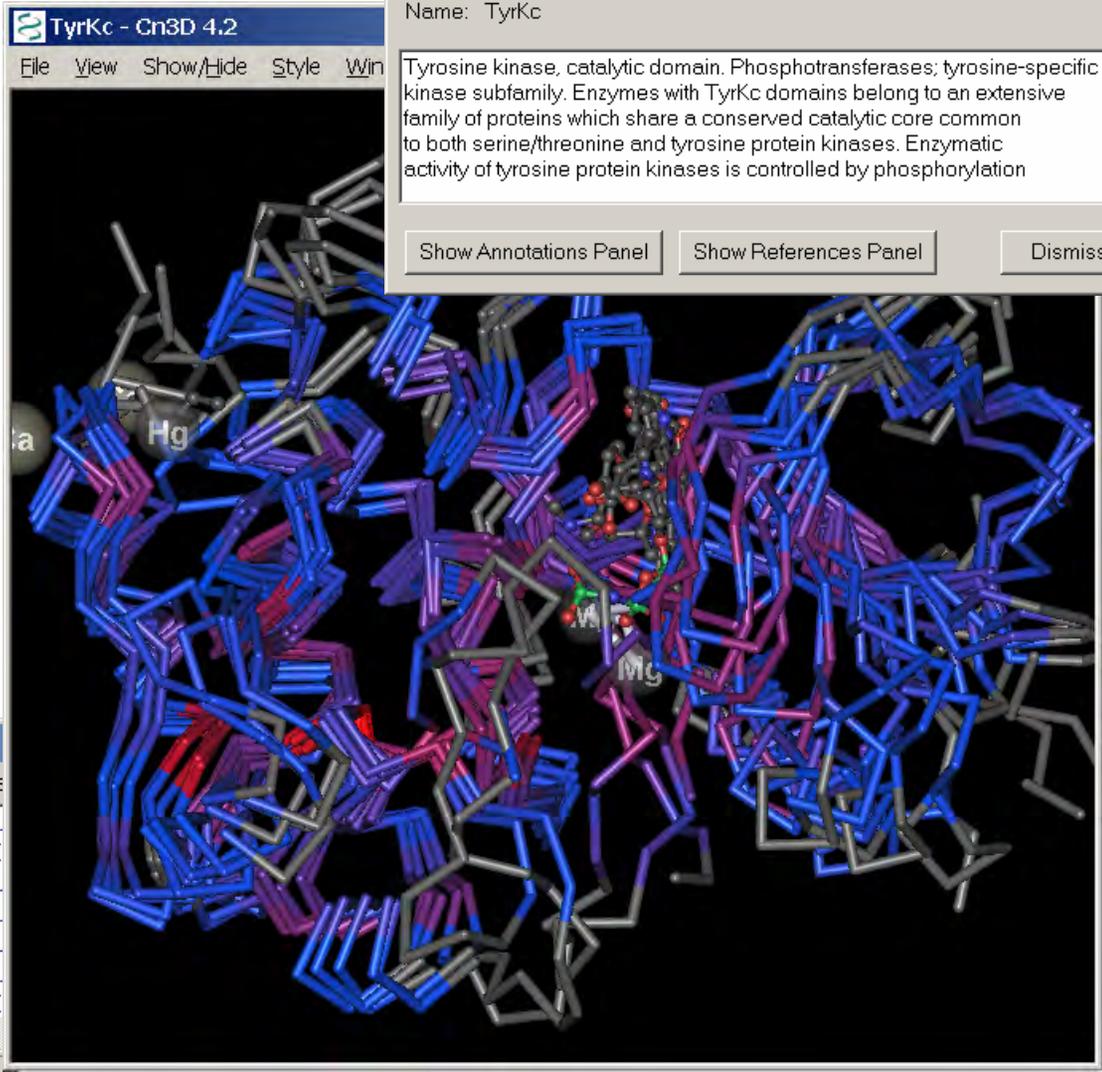


2HCK - Sequence/Alignment Viewer

View Edit Mouse Mode Unaligned Justification Imports

```
2HCK_A msveaflae anvmktlqhdklvklhavvtkepiyiitefmakgsl  
2HCK_B msveaflae anvmktlqhdklvklhavvtkepiyiitefmakgsl
```

# Tyrosine Kinase Family Member ...



**CDD Descriptive Items**

Name: TyrKc

Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-specific kinase subfamily. Enzymes with TyrKc domains belong to an extensive family of proteins which share a conserved catalytic core common to both serine/threonine and tyrosine protein kinases. Enzymatic activity of tyrosine protein kinases is controlled by phosphorylation

Show Annotations Panel Show References Panel Dismiss

**TyrKc - Sequence/Alignment Viewer**

View Edit Mouse Mode Unaligned Justification Imports

<i>IFGI_A</i>	l p e d p r W E L P R D R L V L g ~ ~ ~ k P I
<i>IBYG_A</i>	e f y r s g W A L N M K E L K L l ~ ~ ~ q T I
<i>IIR3_A</i>	v f v p d e W E V S R E K I T L l ~ ~ ~ r E L
<i>IIRK</i>	v f v p d e W E V S R E K I T L l ~ ~ ~ r E L
<i>QHCK_A</i>	p w e k d a W E I P R E S L K L e ~ ~ ~ k K L
<i>gi 136238</i>	k p d t y v Q H I K R R D I V L k ~ ~ ~ r E L
<i>gi 422541</i>	n a k l l s L E Y P R N N I E Y v ~ ~ ~ r D I
<i>gi 232109</i>	

# Links from “Quercetin” to PubMed ...

PubChem Compound Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Copy Paste

Address <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=480> Go Links

MeSH Synonyms Properties Descriptors Category Exports

**Medical Subject Annotations:** (Total:2) [Display: Next 1 | All](#)

**Quercetin**  
A flavonol widely distributed in plants. It is an antioxidant, like many other phenolic heterocyclic compounds. Glycosylated forms include RUTIN and quercetrin.

[Show MeSH Tree Structure](#)

**PubMed via MeSH** Choose by Subheadings:

<a href="#">administration and dosage</a>	<a href="#">adverse effects</a>	<a href="#">analogs and derivatives</a>
<a href="#">analysis</a>	<a href="#">antagonists and inhibitors</a>	<a href="#">biosynthesis</a>
<a href="#">blood</a>	<a href="#">chemical synthesis</a>	<a href="#">chemistry</a>
<a href="#">diagnostic use</a>	<a href="#">genetics</a>	<a href="#">immunology</a>
<a href="#">isolation and purification</a>	<a href="#">metabolism</a>	<a href="#">pharmacokinetics</a>
<a href="#">pharmacology</a>	<a href="#">physiology</a>	<a href="#">radiation effects</a>
<a href="#">secretion</a>	<a href="#">standards</a>	<a href="#">therapeutic use</a>
<a href="#">toxicity</a>	<a href="#">urine</a>	

**Depositor-Supplied Synonyms:** (Total: 74) [?](#)

Internet

# Links from Compounds to PubMed ...

Entrez PubMed - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Refresh Stop

Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pubmed&term=%22Quercetin%2fantagonists%20and%20> Go Links »

PubMed Services  
Journals Database  
MeSH Database  
Single Citation Matcher  
Batch Citation Matcher  
Clinical Queries  
Special Queries  
LinkOut  
My NCBI (Cubby)

Related Resources  
Order Documents  
NLM Mobile  
NLM Catalog  
NLM Gateway  
TOXNET  
Consumer Health  
Clinical Alerts  
ClinicalTrials.gov  
PubMed Central

2: [Lin HC, Cheng TH, Chen YC, Juan SH](#) Related Articles, Links  
**Mechanism of heme oxygenase-1 gene induction by quercetin in rat aortic smooth muscle cells.**  
Pharmacology. 2004 Jun;71(2):107-12.  
PMID: 15118350 [PubMed - indexed for MEDLINE]

3: [Ko WC, Chen MC, Wang SH, Lai YH, Chen JH, Lin CN](#) Related Articles, Links  
**3-O-methylquercetin more selectively inhibits phosphodiesterase subtype 3.**  
Planta Med. 2003 Apr;69(4):310-5.  
PMID: 12709896 [PubMed - indexed for MEDLINE]

4: [Reuelta MP, Cantabrana B, Hidalgo A](#) Related Articles, Links  
**Mechanisms involved in kaempferol-induced relaxation in rat uterine smooth muscle.**  
Life Sci. 2000 Jun 8;67(3):251-9.  
PMID: 10983869 [PubMed - indexed for MEDLINE]

5: [Ciolino HP, Daschner PJ, Yeh GC](#) Related Articles, Links  
**Dietary flavonols quercetin and kaempferol are ligands of the aryl hydrocarbon receptor that affect CYP1A1 transcription differentially.**  
Biochem J. 1999 Jun 15;340 (Pt 3):715-22.  
PMID: 10359656 [PubMed - indexed for MEDLINE]

6: [Lagarrigue S, Chaumontet C, Heberden C, Martel P, Gaillard-Sanchez I](#) Related Articles, Links  
**Suppression of oncogene-induced transformation by quercetin and retinoic acid in rat liver epithelial cells.**  
Cell Mol Biol Res. 1995;41(6):551-60.  
PMID: 8855241 [PubMed - indexed for MEDLINE]

Internet

# Link to BioAssays where “Active” ...

PubChem BioAssay - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Go Links

Address: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pccompound&cmd=Display&dopt=pccompound\\_pcassay\\_...](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pccompound&cmd=Display&dopt=pccompound_pcassay_...)

Items 1 - 9 of 9 One page.

<input type="checkbox"/> 1: AID: <a href="#">396</a>	Compound Screen Assay, Human YWHAB Source: SGC0xCompounds Total substances tested:1; Active:1	Links
<input type="checkbox"/> 2: AID: <a href="#">395</a>	Compound Screen Assay, Human STK16 Source: SGC0xCompounds Total substances tested:35; Active:35	Links
<input type="checkbox"/> 3: AID: <a href="#">393</a>	Compound Screen Assay, Human PIM1 Source: SGC0xCompounds Total substances tested:91; Active:91	Links
<input type="checkbox"/> 4: AID: <a href="#">386</a>	Compound Screen Assay, Human DIRAS Source: SGC0xCompounds Total substances tested:2; Active:2	Links
<input type="checkbox"/> 5: AID: <a href="#">383</a>	Compound Screen Assay, Human CLK3	Links

PubChem  
Help | FAQ

PubChem  
Substance  
Structures supplied by depositors

PubChem  
Compound  
Unique structures with computed properties

PubChem  
BioAssay  
Bioactivity assay results supplied by depositors

PubChem

Local intranet

# BioAssay where “Active” ...

AID 367 - PubChem BioAssay Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=367&c> Go

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

### Select BioAssay Results

Test Results:

BioAssay Display Option

Group Results By:

Duplicate Test Option:

AID: 367

**Name:** HIV-2 RNase H Inhibition  
**Data Source:** [MTDP](#) (RNAI)  
**BioAssay Version:** 1.2

Other Filters:

Substance Filter

Internet

# BioAssay where “Active” ...

AID 367 - PubChem BioAssay Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=367&c> Go

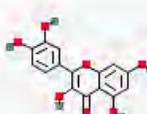
**AID: 367**  
**Name:** HIV-2 RNase H Inhibition  
**Data Source:** [MTDP](#) (RNAI)  
**BioAssay Version:** 1.2

Test Results: [Re-Show](#) [Select](#) [Summary](#)

[Results](#) [Download](#)

BioAssay Display Option  
**Group Results By:**   
**Duplicate Test Option:**

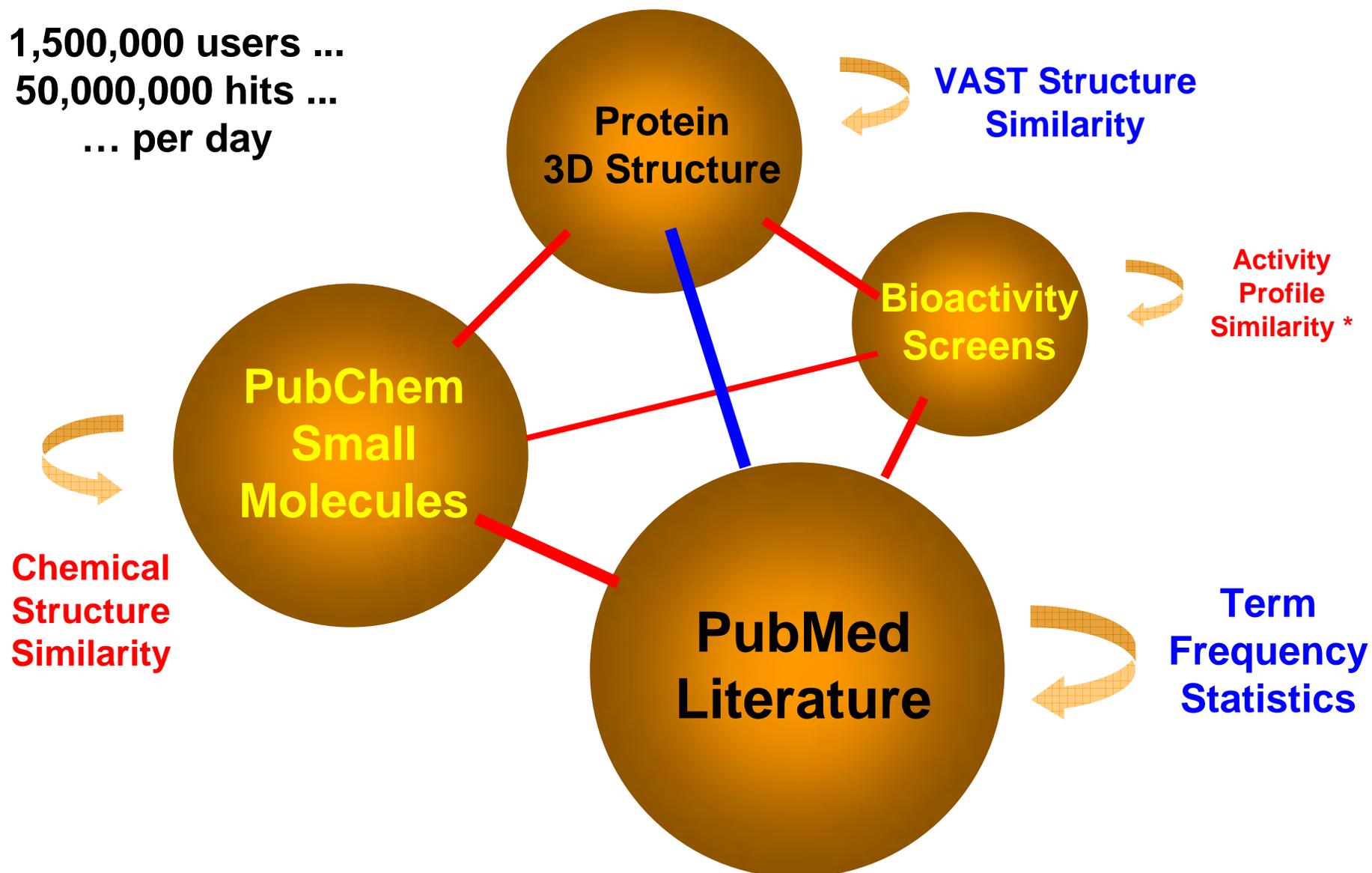
Selected BioAssay Result Count: 1 out of 206 Page: 1 of 1

#	Structure	CID	Score	Outcome	Links	IC-50 UM
1		<a href="#">5280343</a>	26	Active	9	

Done Internet

# Entrez Links and Neighbors ...

1,500,000 users ...  
50,000,000 hits ...  
... per day



# PubChem Retrieval System ...

... Optimize “discoverability” for molecular biologists by integrating PubChem into NCBI’s Entrez / PubMed Search Engine

... Chemical structure search

... Bioassay result search

... Exploratory structure-activity tools

# Quercetin Structure-Activity ...

PubChem Compound - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Select+1> Go

NCBI PubChem Compound National Library of Medicine NLM My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure PMC PubChem Books

Search PubChem Compound for #4 #11 Go Clear Save Search

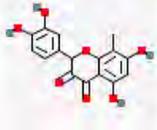
Limits Preview/Index History Clipboard Details

Display PubChem BioActivity Summary Show 20 Sort by Send to

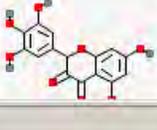
All: 49 BioAssay: 49 Protein3D: 4 Rule of 5: 44

Items 1 - 20 of 49 Page 1 of 3 Next

▮ 1: CID: [6419774](#) Related Structures, Assays

  
CNC-244128, S00072  
IUPAC: 2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-8-methyl-chroman-3,4-dione  
MW: 316.262 | MF: C16H12O7

▮ 2: CID: [6419772](#) Related Structures, Assays

  
CNC-364524, S00070  
IUPAC: 5,7-dihydroxy-2-(3,4,5-trihydroxyphenyl)chroman-3,4-dione  
MW: 318.235 | MF: C15H10O8

Local intranet

# PubChem Bioactivity Summary ...

PubChem BioAssay Summary - Microsoft Internet Explorer

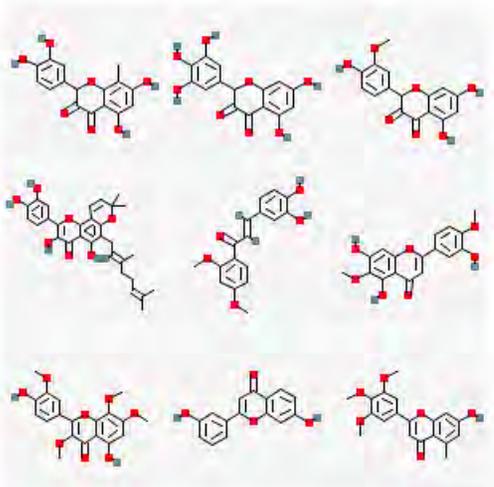
File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?q=cids&qu> Go

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

## BioActivity Summary:



-  **Total Compound count: 49**  
(9 structures shown)  
Total tested: 49  
Active: 49  
Inactive: 43
-  **Structure Activity Analysis**
-  **Structure Clustering**
-  **Selected BioAssays to Entrez**

Total BioAssays: 193

javascript:AssayQueryHandler('heat'); Local intranet

# PubChem Bioactivity Summary ...

PubChem BioAssay Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?q=cids&qui> Go

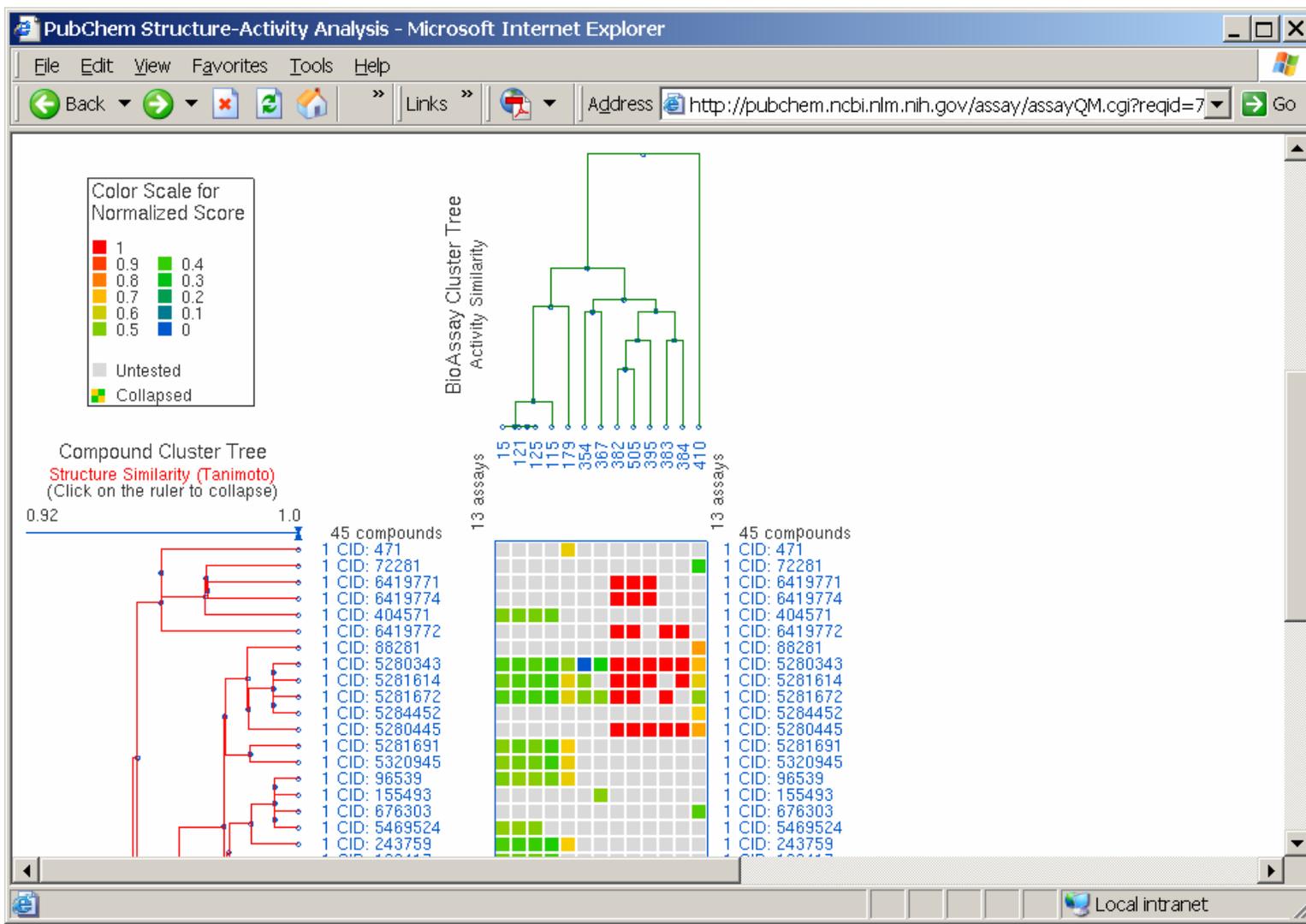
Total BioAssays: 193  
Total Pages: : 2

Display: 100 Go To Page 1

#	<input type="checkbox"/> AID	Active	Inactive	Total Tested	Data	Name
1	<input checked="" type="checkbox"/>	410	22	1	23	Data p450-cyp1a2
2	<input checked="" type="checkbox"/>	382	11		11	Data Compound Screen Assay, Human CLK1
3	<input checked="" type="checkbox"/>	505	9		9	Data Compound Screen Assay, Human PIM2
4	<input checked="" type="checkbox"/>	395	8		8	Data Compound Screen Assay, Human STK16
5	<input checked="" type="checkbox"/>	354	7	2	9	Data NCTR Estrogen Receptor Binding Database (NCTRER)
6	<input checked="" type="checkbox"/>	383	5		5	Data Compound Screen Assay, Human CLK3
7	<input checked="" type="checkbox"/>	115	5	12	17	Data NCI human tumor cell line growth inhibition assay. Data for the SR Leukemia cell line
8	<input checked="" type="checkbox"/>	384	4		4	Data Compound Screen Assay, Human CSNK1G2
9	<input checked="" type="checkbox"/>	367	4		4	Data HIV-2 RNase H Inhibition
10	<input checked="" type="checkbox"/>	179	4	15	19	Data NCI AIDS Antiviral Assay
11	<input checked="" type="checkbox"/>	125	4	16	20	Data NCI human tumor cell line growth inhibition assay. Data for the HL-60(TB) Leukemia cell line

Local intranet

# PubChem Structure-Activity ...



# Active Compound Cluster ...

PubChem Compound - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Refresh Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?WebEnvRq=1> Go

About Entrez  
Entrez Help

PubChem  
Help | FAQ

PubChem  
Substance  
Structures  
supplied by  
depositors

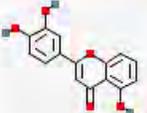
PubChem  
Compound  
Unique structures  
with computed  
properties

PubChem  
BioAssay  
Bioactivity assay  
results supplied by  
depositors

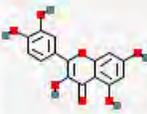
All: 6 BioAssay: 6 Protein3D: 3 Rule of 5: 4

Items 1 - 6 of 6 One page.

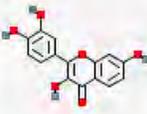
1: CID: [88281](#) Related Structures, Assays

  
TNP00056, ZINC00039317 ...  
IUPAC: 2-(3,4-dihydroxyphenyl)-5-hydroxy-chromen-4-one  
MW: 270.237 | MF: C15H10O5

2: CID: [5280343](#) Related Structures, Assays, Literature, Other Entrez Databases

  
quercetin, Sophoretin ...  
IUPAC: 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-chromen-4-one  
MW: 302.236 | MF: C15H10O7

3: CID: [5281614](#) Related Structures, Assays, Literature, Other Entrez Databases

  
Fisetin, Cotinin ...  
IUPAC: 2-(3,4-dihydroxyphenyl)-3,7-dihydroxy-chromen-4-one  
MW: 286.236 | MF: C15H10O6

4: CID: [5281677](#) Related Structures, Assays, Literature, Other Entrez Databases

Local intranet

# BioAssay Cluster ...

PubChem BioAssay - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?WebEnvRq=1> Go

**PubChem**  
Help | FAQ

PubChem  
Substance  
Structures  
supplied by  
depositors

PubChem  
Compound  
Unique structures  
with computed  
properties

PubChem  
BioAssay  
Bioactivity assay  
results supplied by  
depositors

PubChem  
Structure Search

□ 1: AID: [384](#) Links  
Compound Screen Assay, Human CSNK1G2  
Source: SGC0xCompounds  
Total substances tested: 29; Active: 29

□ 2: AID: [383](#) Links  
Compound Screen Assay, Human CLK3  
Source: SGC0xCompounds  
Total substances tested: 55; Active: 55

□ 3: AID: [395](#) Links  
Compound Screen Assay, Human STK16  
Source: SGC0xCompounds  
Total substances tested: 55; Active: 55

□ 4: AID: [505](#) Links  
Compound Screen Assay, Human PIM2  
Source: SGC0xCompounds  
Total substances tested: 67; Active: 67

□ 5: AID: [382](#) Links  
Compound Screen Assay, Human CLK1  
Source: SGC0xCompounds  
Total substances tested: 121; Active: 121

<http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=505> Local intranet

# Another BioAssay Cluster ...

PubChem BioAssay - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?WebEnvRq=1> Go

About Entrez  
Entrez Help

PubChem  
Help | FAQ

PubChem  
Substance  
Structures  
supplied by  
depositors

PubChem  
Compound  
Unique structures  
with computed  
properties

PubChem  
BioAssay  
Bioactivity assay  
results supplied by  
depositors

All: 4

Items 1 - 4 of 4 One page.

1: AID: [115](#) Links  
NCI human tumor cell line growth inhibition assay. Data for the SR Leukemia cell line  
Source: DTP/NCI  
Total substances tested: 35247; Active: 3567

2: AID: [125](#) Links  
NCI human tumor cell line growth inhibition assay. Data for the HL-60(TB) Leukemia cell line  
Source: DTP/NCI  
Total substances tested: 38933; Active: 3681

3: AID: [121](#) Links  
NCI human tumor cell line growth inhibition assay. Data for the K-562 Leukemia cell line  
Source: DTP/NCI  
Total substances tested: 41721; Active: 3228

4: AID: [15](#) Links  
NCI human tumor cell line growth inhibition assay. Data for the NCI-H522 Non-Small Cell Lung cell line  
Source: DTP/NCI  
Total substances tested: 38588; Active: 3026

Local intranet

# PubMed Connection ...

Entrez PubMed - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Displayf> Go

NCBI PubMed A service of the National Library of Medicine and the National Institutes of Health [My NCBI](#) [\[Sign In\]](#) [\[Register\]](#)

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search PubMed for leukemia kinase inhibitor Go Clear [Save Search](#)

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

All: 3124 Review: 281

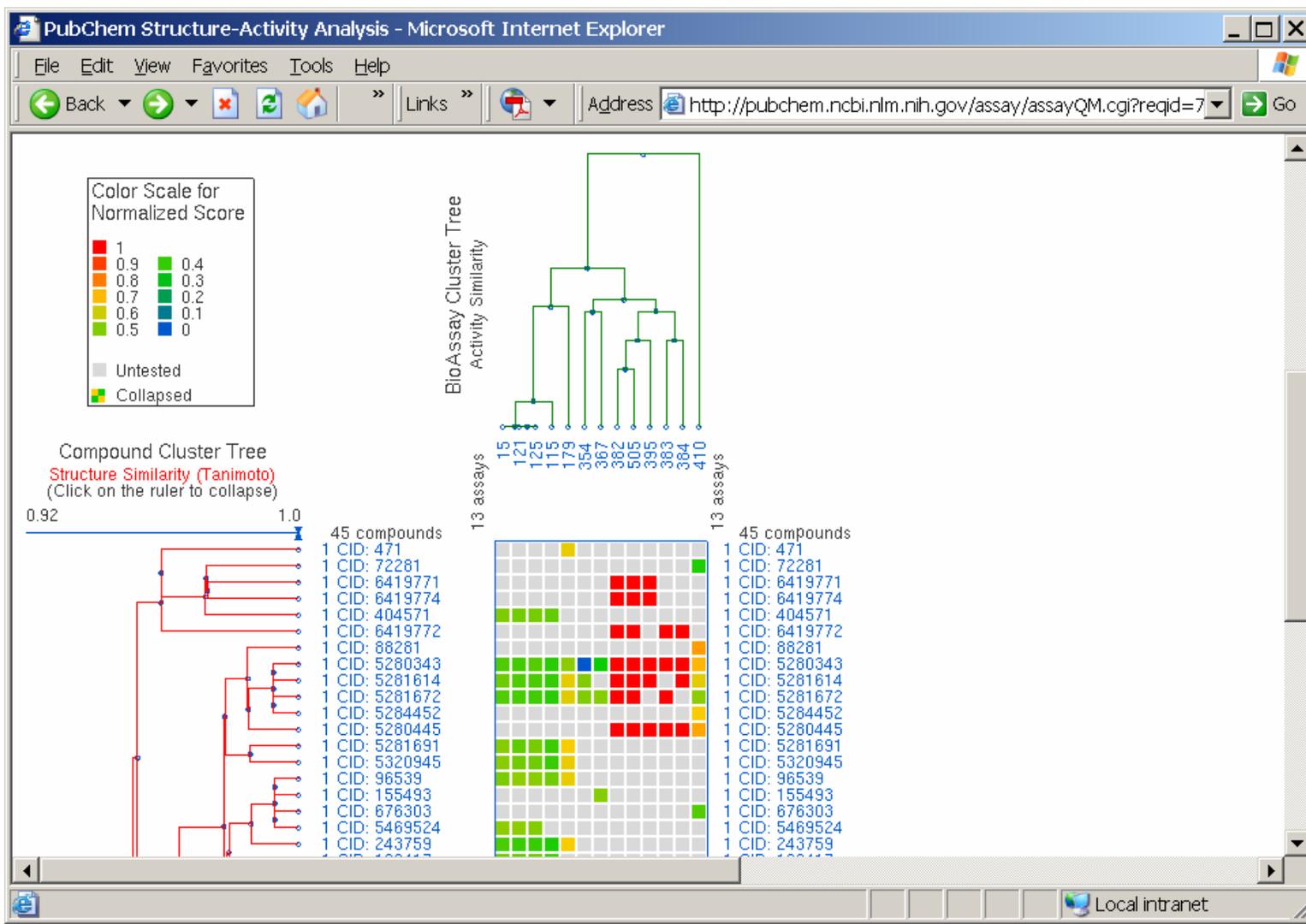
Items 1 - 20 of 281 Page 1 of 15 Next

1: [Larsen TS, Pallisgaard N, Christensen JH, Gram-Hansen P, Kerndrup GB, Moller MB, Hasselbalch HC.](#) Related Articles, Links  
**[New molecular markers within the chronic myeloproliferative disorders. II: the JAK2 mutation]**  
Ugeskr Laeger. 2006 Sep 25;168(39):3299-303. Review. Danish.  
PMID: 17032592 [PubMed - indexed for MEDLINE]

2: [Witzig TE, Kaufmann SH.](#) Related Articles, Links  
**Inhibition of the phosphatidylinositol 3-kinase/mammalian target of rapamycin pathway in hematologic malignancies.**  
Curr Treat Options Oncol. 2006 Jul;7(4):285-94. Review.  
PMID: 16916489 [PubMed - indexed for MEDLINE]

Local intranet

# PubChem Structure-Activity ...

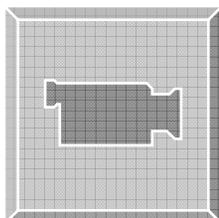


# PubChem BioAssay Tools Demo ...

The screenshot shows a web browser window titled "PubChem BioAssay - Windows Internet Explorer". The address bar contains the URL `http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pcassay`. The search bar shows the query `PubChem BioAssay` for `[679[uid] OR 706[uid]]`. The search results are displayed in a list format, showing two items:

- Item 1:** AID: [706](#)  
Yeast Lifespan Shortening Chemical Screening, Restrictive Growth Control - Pilot Screen  
Source: SRMLSC  
Total substances tested:9532; Active:75
- Item 2:** AID: [679](#)  
Factor XIa Dose Response Confirmation  
Source: PCMD  
Total substances tested:120; Active:30

The interface includes navigation tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "PMC", "PubChem", and "Books". A sidebar on the left contains links for "About Entrez", "Entrez Help", "PubChem Help | FAQ", "PubChem Substance Structures supplied by depositors", and "PubChem". The bottom status bar shows "Internet" and "100%" zoom.



# Yeast Replicative Lifespan Assay ...

The screenshot shows a Windows Internet Explorer browser window displaying the BioAssay Summary for AID: 706. The address bar shows the URL <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=706>. The page title is "BioAssay Summary".

**BioAssay Summary**

**AID: 706**  
**Name:** Yeast Lifespan Shortening Chemical Screening, Restrictive Growth Control - Pilot Screen  
**Data Source:** SRMLSC (Yeast Lifespan Pilot - restrictive)

**Test Results:** [Show](#) [Select](#) [Plot](#)

**Structure Activity Analysis:** [Show](#)

**Structure Clustering:** [Show](#)

**Neighbors, Related BioAssays:**

- Activity Overlap:** [81](#) [Summary](#)
- Compound BioActivity Summary:** [Active](#) [All](#)
- Substance BioActivity Summary:** [Active](#) [All](#)

The browser's taskbar at the bottom shows the "Internet" icon and a zoom level of 100%.

# BioAssay Description ...

**Description:**

Southern Research Molecular Libraries Screening Center (SRMLSC)  
Southern Research Institute (Birmingham, Alabama)  
NIH Molecular Libraries Screening Centers Network (MLSCN)  
Assay Provider: Dr. David S. Goldfarb, University of Rochester  
Award: R03 MH076395-01

There is now solid evidence for the early evolution of conserved pathways for aging. These pathways may allow eukaryotic cells and animals to postpone reproduction in unfavorable environmental conditions. Key elements of public cellular mechanisms that extend or shorten lifespan, including the role of sirtuins in lifespan, were first discovered in *Saccharomyces cerevisiae*, which provides a strong rationale for the use of yeast as a model genetic system for studying aging. For example, lifespan extension by caloric restriction occurs in both yeast and rodents.

Yeast replicative lifespan is the number of times a mother cell replicates before she senesces and dies. The replicative lifespan of a yeast strain is described by the mean or median lifespan of a cohort of mother cells, which can vary widely among laboratory strains, but is normally between 20-25 generations. The clock for daughters is generally reset to zero although daughters of older mothers, which replicate more slowly, have reduced lifespans. The genetic program(s) that sets the clock, and the cellular mechanisms that respond to environmental cues to extend lifespan, such as caloric restriction, are poorly understood.

Here we use a genetically modified version of a high throughput replicative lifespan assay called the DeaD assay (1). Under permissive conditions, in galactose-containing medium, the culture divides exponentially. Under restrictive conditions, in glucose medium, the daughters show a great propensity to die, and the saturation point of the culture is limited by the lifespan of the of mother cells rather than nutrient limitation. Compounds that

# PubMed search for “Sirtuins” ...

The screenshot shows a Windows Internet Explorer browser window displaying the PubMed website. The address bar shows the URL: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=>. The page header includes the NCBI logo, the PubMed logo, and the text "A service of the National Library of Medicine and the National Institutes of Health". There are links for "My NCBI", "Sign In", and "Register".

The search interface shows "Search PubMed for" with a search box and "Go" and "Clear" buttons. Below the search box are buttons for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The "Display" section shows "AbstractPlus" selected, "Show 20", "Sort by", and "Send to". There are also buttons for "All: 1" and "Review: 1".

The search results show one result: "1: [Trends Mol Med.](#) 2007 Feb;13(2):64-71. Epub 2007 Jan 4." There is a link to the "FULL-TEXT ARTICLE" and a "Links" button.

The abstract text is: **SIR2: a potential target for calorie restriction mimetics.** [Chen D, Guarente L.](#) Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. Calorie restriction (CR) extends lifespan in a wide variety of species and mitigates diseases of aging in mammals. Here, we describe the evidence that the silent information regulator 2 (SIR2) gene, which encodes a nicotinamide adenine dinucleotide (NAD)-dependent deacetylase, regulates lifespan and mediates CR in lower species such as *Saccharomyces cerevisiae* and *Caenorhabditis elegans*. We discuss the emerging role of mammalian SIR2 homologs in regulating

On the right side, there is a "Related Links" section with the following items:

- ▶ Sirtuin-independent effects of nicotinamide on lifespan [Aging Cell. 2006]
- ▶ HST2 mediates SIR2-independent lifespan extension by calorie [Science. 2005]
- ▶ Small molecule activators of sirtuins extend *Saccharomyces c* [Nature. 2003]
- ▶ Sirtuin activators mimic caloric restriction and delay age [Nature. 2004]
- ▶ Nicotinamide and PNC1 govern lifespan extension by calorie rest. [Nature. 2003]

The browser status bar at the bottom shows "Internet" and "100%".

# Bioactivity Summary for Actives ...

PubChem BioAssay Summary - Windows Internet Explorer

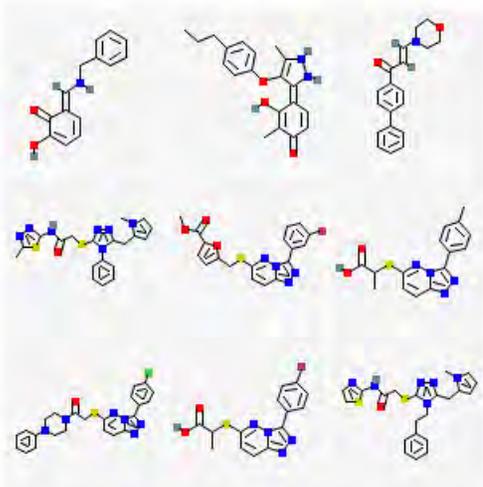
http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?q=cids&query\_key=88&qstatus=c

SID 24268626... PubChem Bi... Entrez PubMed

NCBI PubChem National Library of Medicine NLM

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

### BioActivity Summary:



-  **Total Compound count: 75**   
(9 structures shown)  
Total tested: 75  
Active: 75  
Inactive: 75
-  **Revise Compound Selection:** 
  - Select Active
  - Add Active
  - Add Tested
  - Add Similar Compounds
-  **Structure Activity Analysis** 
-  **Structure Clustering** 
-  **Selected BioAssays to Entrez** 

http://www.ncbi.nlm.nih.gov/ Internet 100%

# Bioactivity Summary for Actives ...

PubChem BioAssay Summary - Windows Internet Explorer

http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?q=cids&query\_key=88&qstatus=c

SID 24268626... PubChem Bi... Entrez PubMed

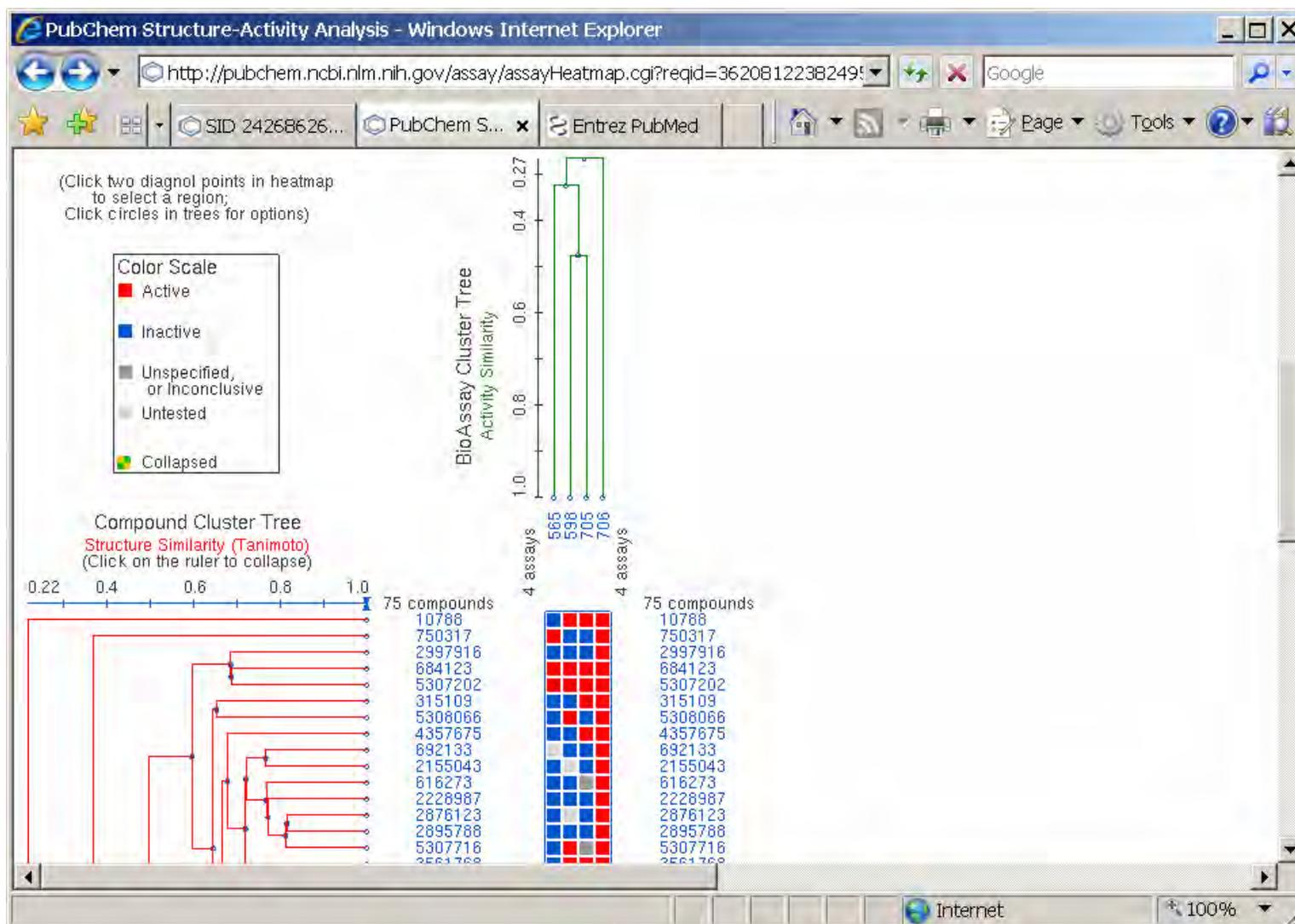
Total BioAssays: 203  
Total Pages: : 3

Display: 100 Go To Page 1

#	<input type="checkbox"/>	AID	Active	Inactive	Total Tested	Data	Name
1	<input checked="" type="checkbox"/>	706	75		75	Data	Yeast Lifespan Shortening Chemical Screening, Restrictive Growth Control - Pilot Screen
2	<input checked="" type="checkbox"/>	705	22	48	75	Data	Yeast Lifespan Shortening Chemical Screening, Permissive Growth Control - Pilot Screen
3	<input checked="" type="checkbox"/>	565	20	54	74	Data	HIV-1 RT-RNase H MLSCN HTS MH077605
4	<input checked="" type="checkbox"/>	598	19	47	66	Data	Human H69AR Lung Tumor Cell Growth Inhibition Assay - 86K Screen
5	<input type="checkbox"/>	641	11	59	70	Data	Allosteric Modulators of D1 Receptors: Primary Screen
6	<input type="checkbox"/>	642	8		8	Data	Allosteric Modulators of D1 Receptors: Confirmation Screen
7	<input type="checkbox"/>	647	8		8	Data	Allosteric Modulators of D1 Receptors: Secondary Assay 2
8	<input type="checkbox"/>	571	8	64	72	Data	Primary Cell Based High Throughput Screening Assay for Antagonists of the 5-

Internet 100%

# Structure-Activity Analysis ...



# Growth Inhibition Assays Cluster ...

The screenshot shows a web browser window titled "PubChem BioAssay - Windows Internet Explorer". The address bar contains the URL: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Select+from+History&db=pca>. The page header includes the NCBI logo, the PubChem BioAssay logo, and the National Library of Medicine (NLM) logo. There are links for "My NCBI", "Sign In", and "Register".

The navigation bar includes tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "PMC", "PubChem", and "Books". The search bar contains the text "PubChem BioAssay" and "for". There are "Go" and "Clear" buttons.

Below the search bar, there are buttons for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The "Display" dropdown is set to "Summary", and the "Show" dropdown is set to "20". There is a "Send to" dropdown.

The search results show "All: 2", "MLSCN: 2", and "Protein Target: 0". It indicates "Items 1 - 2 of 2" and "One page.".

The first result is:

- 1: AID: [705](#) Links  
Yeast Lifespan Shortening Chemical Screening. Permissive Growth Control - Pilot Screen  
Source: SRMLSC  
Total substances tested:9532; Active:23

The second result is:

- 2: AID: [598](#) Links  
Human H69AR Lung Tumor Cell Growth Inhibition Assay - 86K Screen  
Source: SRMLSC  
Total substances tested:85210; Active:5142

The left sidebar contains links for "About Entrez", "Entrez Help", "PubChem Help | FAQ", "PubChem Substance Structures supplied by depositors", and "PubChem". The status bar at the bottom shows "Done", "Internet", and "100%".

# Specific Replicative Lifespan Probes ?

PubChem Compound - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Select+from+History&db=pcc

PubChem Compound

All Databases PubMed Nucleotide Protein Genome Structure PMC PubChem Books

Search PubChem Compound for [ ] Go Clear

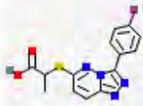
Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

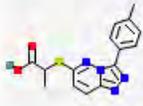
All: 2 BioAssay: 2 Protein3D: 0 Rule of 5: 2

Items 1 - 2 of 2 One page.

1: CID: [5309980](#) Related Structures, Assays

  
MLS000100414, SMR000016796  
IUPAC: 2-[[9-(4-fluorophenyl)-1,2,7,8-tetraazabicyclo[4.3.0]nona-2,4,6,8-tetraen-3-yl]sulfanyl]propanoic acid  
MW: 318.327 | MF: C14H11FN4O2S

2: CID: [5310097](#) Related Structures, Assays

  
MLS000100415, SMR000016805  
IUPAC: 2-[[9-(4-methylphenyl)-1,2,7,8-tetraazabicyclo[4.3.0]nona-2,4,6,8-tetraen-3-yl]sulfanyl]propanoic acid  
MW: 314.363 | MF: C15H14N4O2S

Internet 100%

# Factor XI Inhibition Assay ...

The screenshot shows a Windows Internet Explorer browser window displaying the PubChem BioAssay Summary for AID 679. The address bar shows the URL: <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=679>. The page title is "BioAssay Summary".

**AID: 679**  
Name: [Factor XIa Dose Response Confirmation](#)  
Data Source: PCMD (FXIA\_DRCONF)

Test Results: [Show](#) [Select](#) [Plot](#)

Structure Activity Analysis: [Show](#)

Structure Clustering: [Show](#)

Neighbors | [Links](#) | [Description](#) | [Protocol](#) | [Comment](#) | [Definitions](#)

Neighbors, Related BioAssays:

- Activity Overlap: [34](#) [Summary](#)
- Target Similarity: [6](#) [Summary](#)
- Compound BioActivity Summary: [Active](#) [All](#)
- Substance BioActivity Summary: [Active](#) [All](#)

The browser's taskbar at the bottom shows the "Internet" icon and a zoom level of 100%.

# Factor XI Inhibition Assay ...

The screenshot shows a Windows Internet Explorer browser window displaying the PubChem BioAssay Summary for AID 679. The browser's address bar shows the URL <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=679>. The page title is "BioAssay Summary".

The main content area displays the following information:

- AID: 679** (with a help icon)
- Name:** [Factor XIa Dose Response Confirmation](#)
- Data Source:** [PCMD \(FXIA\\_DRCONF\)](#)

Below this information are three interactive sections, each with a "Show" button:

- Test Results:** [Show](#) [Select](#) [Plot](#)
- Structure Activity Analysis:** [Show](#)
- Structure Clustering:** [Show](#)

A horizontal navigation bar contains the following tabs: [Neighbors](#), [Links](#), [Description](#), [Protocol](#), [Comment](#), and [Definitions](#).

Under the "Neighbors, Related BioAssays:" section, there are four rows of data:

- Activity Overlap:** 34 [Summary](#)
- Target Similarity:** 6 [Summary](#)
- Compound BioActivity Summary:** [Active](#) [All](#)
- Substance BioActivity Summary:** [Active](#) [All](#)

The browser's status bar at the bottom shows "Internet" and a zoom level of "100%".

# Factor XI Inhibition Assay ...

**Description:**

Molecular Library Screening Centre Network (MLSCN)  
Penn Center for Molecular Discovery (PCMD)  
Assay Provider: Scott L. Diamond, University of Pennsylvania  
MLSCN Grant: X01-MH076406-01

**Target**

Factor XI (FXI) circulates as a complex with high molecular weight kininogen (HK) in the plasma at a concentration of 5 ug/ml (equivalent to 31.3 nM, dimeric concentration) as a homodimeric glycosylated blood plasma zymogen of approximately 160 kDa, containing monomeric subunits of 80 kDa each (1). Thrombin (2, 3), factor XIa (FXIa) (3), and factor XIIa alpha (FXIIa) (4), all cleave an internal R369-I370 bond in each monomer of FXI, yielding the enzyme FXIa. The primary role of FXIa has been recognized to be part of a feedback loop resulting from thrombin activation of FXI after inhibition of the extrinsic pathway by tissue factor pathway inhibitor (TFPI).

After activation from FXI to FXIa, FXIa possesses a heavy chain of 369 residues and a light chain of 238 residues. The heavy chain consists of four apple domains (A1-A4) and the light chain represents a trypsin-like serine protease domain with active site residues at H413, D464, and S557 (1, 5-7). FXIa catalyzes FIX to FIXa activation by cleaving two scissile bonds at R145 and R180 (8). The FIXa generated can catalyze FXa formation on the platelet surface with the active cofactor factor VIIIa (FVIIIa), with FVIIIa increasing the Vmax of FX activation by FIXa by 100,000 fold (9). After activation of sufficient levels of FXa by the consolidation pathway, FXa can go on to form a ternary complex with FVa and prothrombin on the platelet surface, to give sufficient levels of thrombin for activation of fibrinogen to fibrin. Formation of this ternary prothrombinase complex in the presence of phospholipids has been shown to increase the rate of prothrombin to thrombin activation

# Bioactivities of Related Targets ...

Pubchem BioAssay Neighbors - Windows Internet Explorer

http://pubchem.ncbi.nlm.nih.gov/assay/assayHeatmap.cgi?reqid=7799455235327

NCBI PubChem National Library of Medicine NLM

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

## Related BioAssays by Target Similarity

**AID: 679**

**Name:** Factor XIa Dose Response Confirmation

**Data Source:** PCMD

**Compounds:** Active: 30

**Targets:** 180352: coagulation factor XI

- Structure-Activity Analysis
- Structure Clustering
- BioActivity Summary
- Selected BioAssays to Entrez

**Related BioAssays of AID 679, Total Count: 6**

Done Internet 100%

# Bioactivities of Related Targets ...

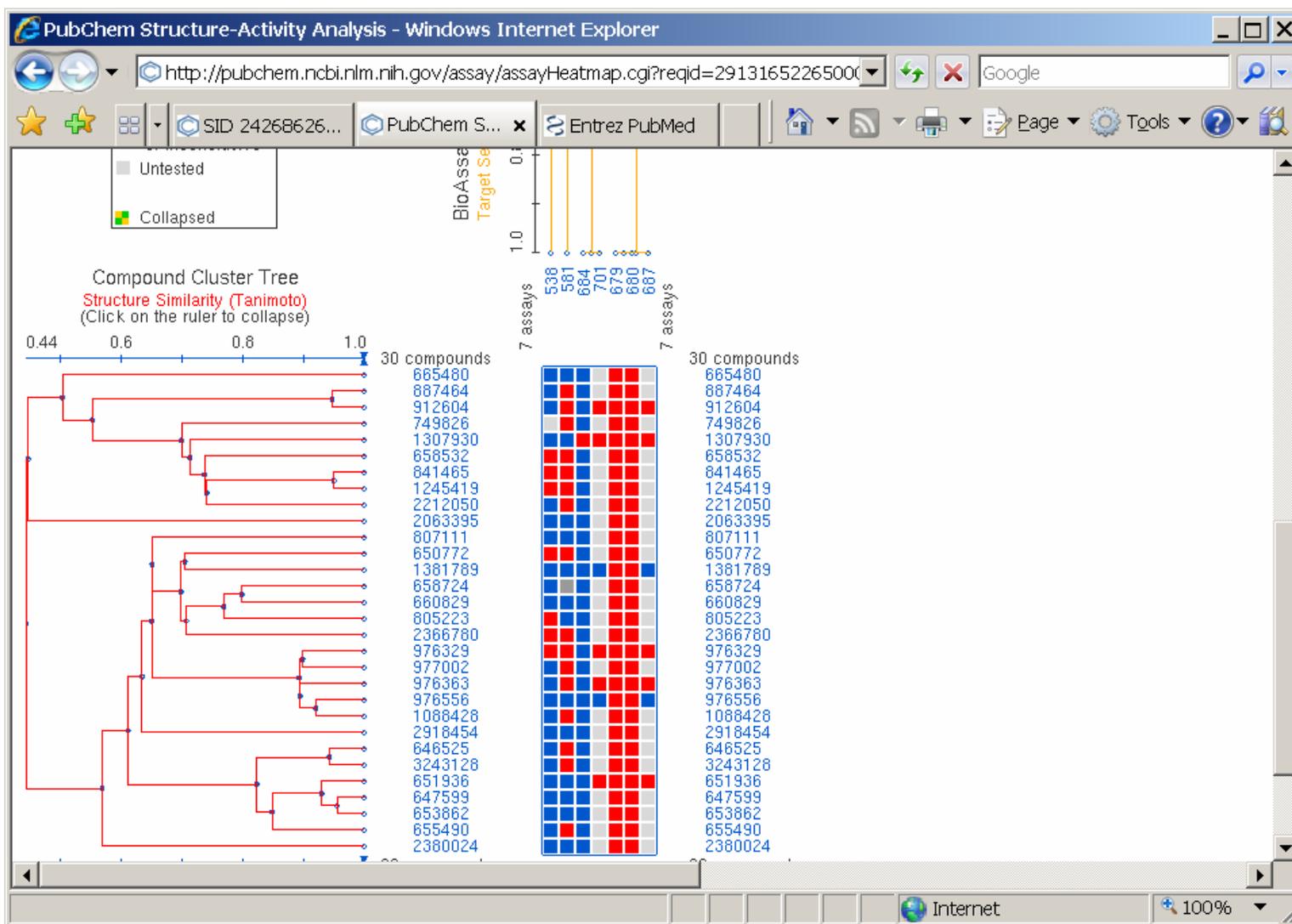
Pubchem BioAssay Neighbors - Windows Internet Explorer

http://pubchem.ncbi.nlm.nih.gov/assay/assayHeatmap.cgi?reqid=7799455235327?

1	<input checked="" type="checkbox"/>	<b>680</b>	1314	0e+00	1	<b>180352</b>	<b>180352</b>	coagulation factor XI
2	<input checked="" type="checkbox"/>	<b>687</b>	1314	0e+00	1	<b>180352</b>	<b>180352</b>	coagulation factor XI Coagulation factor XII precursor (Hageman factor) (HAF) [Contains: Coagulation factor XIIa heavy chain; Beta-factor XIIa part 1; Beta-factor XIIa part 2; Coagulation factor XIIa light chain]
3	<input checked="" type="checkbox"/>	<b>684</b>	155	1e-35	0.362	<b>180352</b>	<b>119763</b>	Coagulation factor XII precursor (Hageman factor) (HAF) [Contains: Coagulation factor XIIa heavy chain; Beta-factor XIIa part 1; Beta-factor XIIa part 2; Coagulation factor XIIa light chain]
4	<input checked="" type="checkbox"/>	<b>701</b>	155	1e-35	0.362	<b>180352</b>	<b>119763</b>	complement component 1, s subcomponent [Homo sapiens]
5	<input checked="" type="checkbox"/>	<b>538</b>	109	1e-21	0.312	<b>180352</b>	<b>4502495</b>	Cathepsin G [Homo sapiens]
6	<input checked="" type="checkbox"/>	<b>581</b>	97	4e-18	0.312	<b>180352</b>	<b>15680217</b>	

Internet 100%

# Structure-Activity Analysis ...



# Specific Factor XI Probes ?

PubChem Compound - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Select+from+History&db=pcc

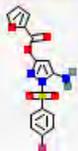
Search PubChem Compound for [ ] Go Clear

Display Summary Show 20 Sort by Send to

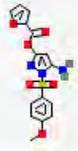
All: 2 BioAssay: 2 Protein3D: 0 Rule of 5: 2

Items 1 - 2 of 2 One page.

1: CID: [647599](#) Related Structures, Assays

 PCMD-CC-MCM-6, ZINC00812583 ...  
IUPAC: [5-amino-1-(4-fluorophenyl)sulfonyl-pyrazol-3-yl] furan-2-carboxylate  
MW: 351.311 | MF: C14H10FN3O5S

2: CID: [653862](#) Related Structures, Assays

 ZINC00812578, BAS 08330403 ...  
IUPAC: [5-amino-1-(4-methoxyphenyl)sulfonyl-pyrazol-3-yl] furan-2-carboxylate  
MW: 363.346 | MF: C15H13N3O6S

Done Internet 100%

<http://pubchem.ncbi.nlm.nih.gov>

**Evan Bolton**

**Jie Chen**

**Svetlana Dracheva**

**Lewis Geer**

**Lianyi Han**

**Jane He**

**Siqian He**

**Wolf-Dietrich Ihlenfeldt**

**Karen Karapetian**

**Vahan Simonyan**

**Ben Shoemaker**

**Wenyao Shi**

**Tugba Suzek**

**Paul Thiessen**

**Valery Tkachenko**

**Jiyao Wang**

**Yanli Wang**

**Jewen Xiao**

**Jian Zhang**

